

A shadow cast on a good cause

Animal rights extremists, like terrorists everywhere, seem to be gathering strength and boldness. Societies that value personal liberty should pay more attention.

FOR many years, there has been a clear need to regularize and make uniform arrangements for safe-guarding the rights of laboratory animals. Luckily, the case for the seemly treatment of captive animals, especially in laboratories, has been widely acknowledged: the general principle of what is fast becoming the legislative norm is that researchers wishing to use laboratory animals for experimental purposes must be licensed to do so, in the process balancing the discomfort and constraints that animals will suffer (given anaesthesia, there is hardly ever pain) against the value of the information they will acquire, often in the presence of government inspectors (in Britain) or laboratory animal care committees (in the United States). It is true that there are some obvious loopholes in the legislation that should be closed: in the United States, for example, regulations promulgated by the National Institutes of Health (NIH) do not formally apply to institutions not receiving federal grants. It is fair to say that much of the improvement brought about in the past decade has arisen because of pressure from lobbying organizations taking animals as their cause.

Now, sadly, the pendulum seems to be swinging too far. The problem now is not the animals and their treatment in captivity, but those who purport to mount self-styled campaigns (the modern idiom) on their behalf. Indeed, there is even a danger that the moderate groups that now share the credit for the improved legislation in force or coming into force will be discredited by the activists, whose most obvious attribute is not their love of animals but their capacity for their hatred of people.

How else to account for the bomb that exploded in a staff common room at the University of Bristol in February, a long way from the nearest animal house? For the threats against the lives of prominent physiologists (not to mention the practice of daubing their houses with paint and other less tolerable substances)? Or for the theft of animals from the INSERM unit at Lyons a few weeks ago (see *Nature* 339, 407; 8 June 1989) whose consequences include the damage done to research programmes nurtured over years and, probably, distress for the stolen animals? The extremists in the animal rights movement are extremists among extremists, fairly to be likened to those who terrorize innocent passengers on hijacked aircraft for obscure causes.

Like some other kinds of terrorists — Greenpeace in its looniest enterprises for example — the animal extremists

derive a measure of disguise from their roots in good-cause organizations of various kinds, some of them highly and deservedly respected. Yet, by now, there is ample evidence that they are sustained not only by their individual incoherent rage at the world in which they find themselves, but by information networks that seem to be increasingly effective at choosing what may be thought suitably vulnerable targets and people. It is chilling that the INSERM laboratory at Lyons should have been attacked only a few weeks after the research article giving offence should have appeared in *Nature*.

What, then, is to be done about them? In principle, civil authorities such as the police forces have powers that could be more vigorously used. In reality, less than might happen will happen so long as the social divisiveness of animal extremism is not clearly recognized for what it is, as a means by which laboratory people conducting themselves entirely lawfully are terrorized and their emulators discouraged. The issue should command the attention not just of that part of the academic community engaged in research with animals, or of the academic community as a whole, but of all who value the civic virtues of personal liberty. Sadly, too little attention has been paid to this shadow on an otherwise good cause. □

Fingerprinting trials

Confusions about the forensic applications of DNA fingerprinting should not disguise the technique's promise.

THE important article by Dr Eric Lander on page 501 of this issue should not be misconstrued. Lander describes pitfalls in the forensic use of the technique of DNA fingerprinting illustrated for him by a judicial prelude to a murder trial in New York City. The court has not yet decided whether evidence of identification based on restriction-fragment analysis of DNA from the victims and the accused will be admissible at the eventual trial, when there will also no doubt be other evidence to consider. But it seems to be accepted that the particular evidence presented of identification by restriction-fragment length is flawed. So does that mean that the DNA techniques have no place in forensic science? Not at all. DNA fingerprinting can be every bit as valuable as the enthusiasts have been saying. The lesson to be drawn



from Lander's tale is that the new technique must be used with care. When has it ever been otherwise with a new technology?

Many of the difficulties are those arising whenever people seek to turn a research technique into a routine assay — techniques borrowed from cytological research for the early diagnosis of cervical cancer, for example, or from strength-of-materials engineering for assessing liability to metal fatigue. In research laboratories, where the objective is usually to test original hypotheses about the natural world, people quite properly adjust the characteristics of their experimental systems to provide the most stringent tests of their hypotheses. Routine assays, by contrast, are best standardized, not merely for convenience, but so that experience will assist empirical understanding of their usefulness and limitations. That is bound to take time.

Lander's article points to several respects in which DNA fingerprinting must be applied with caution, among which the most interesting yet teasing is the demonstration (by an examination of blood-bank samples) that polymorphic alleles are not scattered at random through real human populations. That is hardly a surprise; inter-marriage between ethnic, cultural or even social groups is also far from random. So attempts accurately to calculate the chance that two apparently identical samples of DNA are from different individuals require not only dependable laboratory techniques, but also a knowledge of the distribution of the target alleles among the sub-populations from which the people concerned are drawn. This is a potentially daunting obstacle to the calculation of odds sufficiently astronomical to impress trial juries, however great may be its promise for human anthropologists.

How, meanwhile, should the forensic use of DNA fingerprinting be encouraged without the technique being oversold? The first need is that users should not expect too much. In principle, measurement of the lengths of restriction fragments of people's DNA is no different from the use for forensic purposes of more familiar genetic markers — the A, B and O blood groups, for example. Genetic differences between individuals can be established unambiguously, but there is always a chance that an appearance of genetic identity may be a coincidence. The promise of DNA fingerprinting, still unchallenged, is that the chances of accidental genetic matches will be very much smaller. The confusion at the pre-trial hearings in New York centres on more practical questions, most of them about the need for standard laboratory procedures. (Given that people's liberty is at stake, there is also a need for rules on the custody and use of samples.) If the US National Academy of Sciences believes it could help, it should carry out the study being urged on it without waiting for a subsidy.

Meanwhile, there is a need that everybody should be on guard against the seductiveness of large numbers. When the world's population is of the order of 10^{10} , it cannot much matter that the chance of an accidental genetic match is given as one in 10^{12} or one in 10^{15} . In most circum-

stances, a few extra polymorphic sites could always be used to add a few orders of magnitude to the odds. What juries (and the rest of us) need to understand is that there are exceptions to the sense of safety deriving from these very large numbers — the case of identical twins most obviously, but those arising from flagrant departures from Hardy-Weinberg equilibrium will be more taxing. DNA fingerprinting has a valuable contribution to make to forensic science and the administration of justice, but its evidence, like most other evidence, requires interpretation. (The practical application of gene amplification techniques, which in principle offer identification by a single molecule of DNA, will be similarly beset with difficulties.)

The underlying difficulty is that science and the law differ in temperament. Both pretend to discover truth, science by continually narrowing the range of allowable dispute, the law by its supposition that the judicial process can conjure truth from whatever data are available. Courts are not content with declarations that defendants are, say, "60 per cent guilty", nor should they be. □

When to track AIDS

New York City authorities ask whether the time has come to trace the contacts of AIDS patients. It has not.

CURIOUS though it may be, it is also heartening that the illiberal notions prevalent five years ago for dealing with the then impending spread of AIDS appear to have melted away. Much of the explanation lies in the characteristics of the infection as now recognized: the long and variable period before antibodies make their appearance (allowing for diagnosis) and the long and variable incubation period before the onset of the overt disease, plainly argue against equitable, let alone effective, ostracization. So far, even the most seriously affected communities have resisted even tracing the sexual or needle-sharing partners of those with AIDS on the proper grounds that, in the absence of prophylaxis, knowing who is liable to be struck down would be of little use.

It is natural that the Health Commissioners for New York City should have been (last week, at Montreal — see page 496) the first to call for measures to enforce contact-tracing. The scale of New York's AIDS pandemic is daunting, while there is now some hope that AIDS-associated pneumonia can be prevented by drugs. But the time for compulsory contact-tracing has not yet arrived. First, the efficacy of the prophylaxis needs more secure demonstration. Second, the prospect that contacts would be traced is likely more effectively to dissuade people from seeking treatment than with the simpler (and tractable) venereal diseases. But there remains no assurance that contact-tracing for AIDS, with its inevitable loss of confidentiality, could be undertaken without the pointlessly discriminatory reprisals against the infected that even enlightened communities condone. □

Growing pains for Amgen as epoetin wins US approval

- Congress worries over costs
- Patents disputes continue

Washington

Two weeks ago, Amgen Corporation finally capped eight years of research with permission from the Food and Drug Administration to sell epoetin, a genetically-engineered copy of erythropoietin (EPO) likely to revolutionize treatment of people with kidney failure.

But Amgen still needs stamina to complete the transition from biotechnology company to pharmaceutical supplier: it now faces questions from Congress over its prices on top of continuing patent battles and disputes with its business partner.

Epoetin, which Amgen is selling under the trade name Epogen, is a copy of EPO, a protein produced by the kidneys that triggers the production of red blood cells in bone marrow. Lacking EPO, patients with kidney failure commonly suffer from anaemia. But clinical trials show that treating them with epoetin restores red-blood-cell production, making regular blood transfusions unnecessary.

On 1 June, the Food and Drug Administration (FDA) specifically approved the drug for treatment of anaemia in kidney failure, opening use of the drug to some 95,000 people in the United States. But the market for epoetin is likely to go beyond kidney patients to others needing to increase red-blood-cell production. That could include AIDS and cancer patients and people recovering from operations. According to Shearson Lehman Hutton, a stock analysis company, the US market for epoetin could exceed two thousand million dollars by 1993, making it one of the world's best-selling drugs.

Alongside FDA approval, Epogen was given "orphan drug" status under an act that provides special incentives for treatments for diseases or conditions affecting fewer than 200,000 US patients. The incentives include a seven-year period when competitive manufacture of an identical product is prohibited.

These successes for Amgen have helped make Congress worry that the government will have to pay too much for Epogen through the Medicare scheme that covers kidney-dialysis patients.

Congressman Fortney Stark (Democrat, California), chairman of the House of Representatives Ways and Means health subcommittee is concerned that the price of Epogen, at between \$4,000 and \$8,000 for a year's treatment, would grant Amgen a huge return on its investment. But Amgen points to the legal difficulties that still surround it and the gigantic

gamble it took in investing several hundred million dollars in the research and development of a single product.

On the legal front, Amgen is in triple trouble, embroiled in disputes with its competitors Genetics Institute and Chugai Pharmaceuticals, and a separate dispute over its market share with marketing partner, Johnson & Johnson.

Amgen and Genetics Institute both hold EPO patents. Amgen's patent includes the EPO gene and the genetically-engineered cells that make EPO, but not the final product. Genetics Institute holds a patent for "homogenous [HUMAN] EPO" and takes that to mean all human EPO, even though the company obtained its EPO by purifying it from urine.

As genetically-engineered cells are essential to make large quantities of EPO, the situation is perfect for dispute: Genetics Institute has a claim on the final destination and Amgen on the only way of getting there.

In the meantime, Genetics Institute's licensee, Chugai Pharmaceuticals, has succeeded in importing genetically-engineered EPO made in Japan. Amgen tried to block the imports with an appeal to the International Trade Commission. But, although the judge was sympathetic, nothing could be done to prevent offshore production of recombinant EPO.

Amgen's third quarrel, with Johnson & Johnson's Ortho, seems just as difficult, even though it is a battle with a partner. Their disagreement is over what their agreement means. Amgen granted Johnson & Johnson's Ortho the right to the non-kidney dialysis market for cash and advice. But the companies cannot now agree over the boundaries of their market shares and are engaged in arbitration.

All of Amgen's legal troubles are unlikely to destroy the company. At worst, Amgen may be forced into cross-licensing agreements with other companies. But Congressional interest in the prices of the new products of genetic engineering raises new questions. The US government's decision that Genentech's tissue plasminogen activator was too expensive for the Medicare scheme (see *Nature* 332, 577; 14 April 1988) hit hard at Genentech's plans to become a major pharmaceutical producer. But as a major purchaser of drugs, a spokesman for representative Stark says Congress has "no choice" but to question prices.

Alun Anderson

CHINESE TURMOIL

Expressions of outrage

Washington & Tokyo

THE US National Academy of Sciences suspended its cooperative exchange programme with the Chinese Academy of Sciences and other Chinese institutions last week in response to the killing of demonstrators in Tiananmen Square by Chinese troops. A telegram from Frank Press, President of the National Academy of Sciences to Professor Zhou Guangzhao, president of the Chinese Academy of Sciences expressed "outrage and sadness" and said, "We are shocked and dismayed by the action of Chinese government troops against peaceful demonstrators."

In the United States and Japan, the two most popular destinations for Chinese students, governments moved to reassure Chinese students that they would not be compelled to return home while the present troubles continue. The United

24

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Fang Lizhi in better times.

States has approximately 40,000 students from China, Japan has almost 8,000.

China's best-known dissident, astrophysicist Fang Lizhi, and his wife Li Shuxian, were still in the US Embassy in Beijing on June 12 after seeking sanctuary there a week earlier. The Chinese government has issued arrest warrants for both Fang and Li as part of their crack down on protestors. The US government has said that it does not hand over dissidents in danger of being killed, a stand that is likely to precipitate a diplomatic crisis with the Chinese government. According to the *Washington Post*, Fang has also been criticised by Chinese intellectuals for his "melodramatic" decision to seek US protection because it could be used to support government propaganda that the democracy movement was foreign-inspired.

Alun Anderson & David Swinbanks

Imposing total import bans

Washington

RESPONDING to calls from conservation groups, President George Bush last week announced a total ban on the import of elephant ivory to the United States. The move came after a study by the international Ivory Trade Review Group showed that the population of elephants in Africa has dropped from 1.3 million in 1979 to roughly 750,000 two years ago, mostly from illegal poaching. Limited trade in ivory has so far been permitted by the Convention on the International Trade in Endangered Species (CITES).

Attention to the plight of the elephant was stepped up last year, after the US-based African Wildlife Foundation called for a consumer embargo of ivory products to stem the demand for ivory, and deflate its price (see *Nature* 333, 290; 1988). At the time, such an approach was deemed to be the best way to discourage poaching without a total ban on the sale of ivory.

In October, the 102 parties of CITES



Poachers examine a dead elephant before cutting out the valuable ivory tusks.

will meet in Switzerland to vote on an international across-the-board ban on ivory trade. In the meantime, World Wildlife Fund president Russell Train predicts an "elephant holocaust" across Africa if the rest of the world continues to accept ivory imports.

Carol Ezzell

Paris

THE French environment minister, Brice Lalonde, last week announced that France is to operate a total ban on imports of raw or carved ivory. The decision, and that announced by US President George Bush (see above), follow a call two weeks ago by the British environment minister, Lord Caithness, for a multilateral ban on new ivory trade.

Support has also come from the European Commission, which agreed at a meeting last Thursday to ban all imports of raw and worked ivory. The ban will take effect within the next month, but Britain and France will operate an immed-

iate unilateral ban.

Most affected will be African countries, who have so far been able to continue to trade under Annex 2 of CITES, which allows quotas of exports. Indian elephants already come under Annex 1, which requires a total ban.

By allowing African nations to manage their own quotas under Annex 2, CITES has attracted criticism for, unwittingly, 'laundering' illegal ivory trade. In ten years, the population of African elephants has decreased despite the controls. While the legal export quotas from African countries total about 118 tonnes of ivory each year, Hong Kong, Japan and China last year together declared imports over three times greater than this.

The import bans will be an encouragement to African countries seeking a total ban on international ivory trading. But, unless the import ban is supported by Far East countries, where much ivory is carved, it will do little to curb illegal trading.

Peter Coles

MEDICAL RESEARCH

Health service reform worries MRC

London

MEDICAL research and research training in Britain could suffer as a result of the government's planned reform of the National Health Service (NHS), warns the Medical Research Council (MRC) in a written response to the plans.

Tighter control over budgets by general practitioners and self-governing hospitals could lead to the neglect of medical research, which in the short term adds to the costs of medical care, it says. Although the government states in the white paper that it is committed to maintaining the quality of medical education and research, the MRC complains that there is little information about how this commitment will be honoured. It must ensure, says the MRC, that there is no financial advantage in opting out of medical research, and should introduce positive discrimination to encourage NHS staff to become involved with research. The MRC also calls for the appointment of a director at top level rank to take responsibility for identifying research needs and ensuring that advances are translated into practice.

Other MRC concerns are that it will be more difficult to recruit high quality personnel in academic medicine and clinical research if hospitals can determine pay and conditions for staff, and that fragmentation of the NHS will make it more difficult to carry out multicentre clinical trials, research studies and evaluations of new technologies.

Christine McGourty

Academies' global view

Moscow

THE United States and the Soviet Union are to cooperate to tackle environmental problems. Earlier this month, president of the USSR Academy of Sciences Guriy Marchuk and head of the delegation of the US National Academy of Sciences, president of the National Engineering Academy Robert White, signed a memorandum to establish a United Inter-Academy Committee for Global Ecology. The Committee will comprise the two countries' leading experts; hundreds of research institutes will take part in joint studies.

There was no representative for the nuclear sciences or the nuclear industry at the meeting, perhaps surprising in light of the importance of nuclear power. The opinion of US ecologists was voiced by Robert White, once head of the US hydro-meteorological service. He said that to prevent the greenhouse effect, it was necessary to examine ways of energy production which did not involve discharges of carbon dioxide into the atmosphere. Although public opinion in many countries is hostile to nuclear power (see page 499), safe nuclear power stations can be built today and this technology must be considered alongside other alternative sources of energy.

The first decisions on the co-ordinated directions of studies will be discussed at the next Committee session in Washington next week.

Yuri Kanin/Novosti

COMPUTER ESPIONAGE

Hacker dies

Munich

THE body of a 24-year-old West German computer hacker was found, badly burned, in the woods near Braunschweig on 1 June. The hacker, Karl Koch, was one of two West Germans who told authorities of a plot to pass classified information from Western computer systems to Soviet intelligence agents (see *Nature* 338, 108; 1989).

Authorities found "no evidence" of foul play in Koch's death, according to spokesman Hans Helmut Kehr of the Hildesheim Public Prosecutor's Office. Near Koch's charred remains were found a petrol can and a lighter. Kehr called the case a suicide.

In the week's prior to his death Koch was reportedly in a disturbed condition. But Clifford Stoll of the Harvard Smithsonian Center for Astrophysics and others close to the espionage case consider the suicide hypothesis "doubtful".

Indictments by the West German Federal Prosecutor are expected in the hacker espionage case at the end of the summer.

Steven Dickman

JAPANESE POLITICS

A faltering start for Uno

Tokyo

JAPAN'S new prime minister Sosuke Uno, in his first policy speech before the Diet last week, pledged to "take positive initiatives" to solve global environmental problems, such as the greenhouse effect. He also called for greater development of science and technology and academic research and for the steady advance of nuclear power in Japan. But with the Recruit scandal still dominating Japanese politics, Uno may not be in power long enough to see his policies implemented.

Uno has experience in tackling inter-

Japan's foreign minister, reached agreement in Paris with US Secretary of Energy James Watkin for Japan and the United States to promote nuclear power as a "realistic approach" to global environmental problems.

Uno is expected to put forward further Japanese proposals for the solution of global environmental problems at the upcoming seven-nation annual summit in July. He will also describe the organization of Japan's Human Frontier Science programme at next month's economic summit, provided national representatives of the seven participants can reach agreement on the agenda this week in Tokyo.

Last week, Uno also said that he intends to host an international conference on the global environment in Tokyo in September.

But while such international issues may help divert attention from domestic problems, Uno's prime task in the coming months will be to try and restore public faith in the ruling Liberal Democratic Party (LDP) before this summer's elections for the Upper House of the Diet. Support for the LDP has plunged to an all-time low in the wake of the Recruit scandal and Uno's policy speech was dominated by pledges for political reform.

But Uno is already in trouble: last week the *Sunday Mainichi* magazine published revelations about his personal life from a "geisha" who claimed she had been his mistress. Questions were asked in the Diet, but Uno refused public comment. If the LDP loses its majority in the upper house, Uno could be forced to step down.

David Swinbanks

AP

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Premier Uno faces the media.

national science and technology issues. In 1977, as director general of the Science and Technology Agency, he succeeded in concluding an agreement with the United States allowing Japan to reprocess spent nuclear fuel.

In 1983, Uno was appointed head of the Ministry of International Trade and Industry. And a few days before his appointment as prime minister, Uno, then

JAPAN IN SPACE

Another victim for Recruit?

Tokyo

JAPAN'S plans to participate in the building of the US space station have been seriously delayed by the Recruit scandal, the discovery that many members of the ruling Liberal Party had received financial benefits by being offered a chance to buy publicly-quoted shares at favourable prices. Although the scandal has led to the resignation of prime minister Takeshita and his replacement by Mr Sosuke Uno, the opposition parties are still not satisfied that full details have been uncovered.

One consequence is that a motion to approve the US-Japan inter-governmental agreement on the space station has been stalled in the Diet since April, as part of the opposition's boycott of Diet deliberations. As a result, the National Space Development Agency (NASDA) has not been able to begin on its part in the construction phase of the project.

The foreign-affairs committee of the

lower house of the Diet is expected to debate the inter-governmental agreement this week. But it is uncertain whether the motion, which must also be debated in the upper house, will pass before the Diet closes for the summer recess on 22 June. Among other things, opposition parties are objecting to the possible military uses of the station.

Europe and Canada have already entered the construction phase after signing formal participation agreements last September. But even if the motion passes the Diet, Japanese contractors will be on a very tight schedule, according to Tatsuo Yonezawa of the international affairs division of NASDA. Funds for construction have been available in NASDA's budget since 1988 and contractors are itching to get started. But if the motion does not pass, "nobody knows" when construction can begin, Yonezawa says.

David Swinbanks

POST-ELECTION POLAND

Students' organization soft-pedals

London

POLAND'S Independent Student's Association (NZS) has called off all demonstrations and sit-ins until the autumn. Although the state and party authorities agreed, in the April "Round Table" accords, to recognize the NZS as a legal association, the Warsaw Provincial Court refused last month to register it, on the grounds that its statutes demanded the right to strike. This, the judge ruled, was contrary to the 1985 Higher Education Act, and also to the new law on associations, since — she said — students are not employees and so, under existing legislation, have no right to strike.

After several days of sit-ins, strikes and street blockades throughout Poland, the students filed an appeal to the Supreme Court against the Provincial Court decision. The appeal is due to come up in October, conveniently coinciding with the opening of the next academic year. In the meantime, the students are putting considerable trust in Solidarity's landslide in the Parliamentary elections, and in the pledge of Solidarity's leader, Lech Walesa, that Deputies and Senators elected from the Solidarity list would treat the legalization of NZS as "a matter of urgency". Vera Rich

TOXINS

Neurolathyrism remains a threat in India

New Delhi

WHILE the 40,000 victims of the poison gas tragedy at Bhopal, capital of Madhya Pradesh, may hope to be rehabilitated following a settlement with Union Carbide, a far larger number of victims of poisoning must go without help elsewhere in the same state. These are the poor agricultural labourers who suffer from neurolathyrism, a crippling condition of the nervous system, caused by eating the foodgrain *Lathyrus sativus*.

Lathyrus sativus is a hardy crop that needs no irrigation, fertilizers or pesticides. Its seeds are rich in protein (28 per cent) but they also contain a neurotoxin called beta-N-oxalyl aminoalanine. Studies at the National Institute of Nutrition in Hyderabad 20 years ago showed that the toxin can cause permanent paralysis in monkeys.

Several thousand agricultural labourers in Madhya Pradesh have been living for decades on kesari dhal made with *Lathyrus sativus* and given to them by landlords as payment.

Agricultural scientists have tried for more than a decade to breed a toxin-free variety of *Lathyrus sativus* without success. The Agriculture Ministry claims, however, that the incidence of neurolathyrism is declining in Madhya Pradesh as rice and wheat become cheaper than *Lathyrus sativus*.

K.S. Jayaraman

Cambridge first with local law

Boston

SLOWLY but steadily the city of Cambridge, Massachusetts, is moving towards passage of the first local ordinance in the United States governing the care of laboratory animals.

Last week, before a packed crowd of animal advocates and representatives from the research community, Cambridge City Council voted unanimously to write the recommendations of its advisory panel into the form of a local ordinance. That includes the establishment of a city "Commissioner of Laboratory Animals" and is expected to pass the council with little opposition.

The City Council decision to heed the panel's advice has been greeted favourably. Animal rights groups claim the move towards local regulation as a victory, although some advocate even stricter provisions. Most researchers in the city, meanwhile, welcome the new regulations publicly, but privately are relieved that research will not be restricted by more draconian measures.

The council was under a great deal of pressure from animal groups to act on the issue. The three-member panel, which included a local veterinarian and representatives from the scientific community and animal rights groups, gained respect from opposing sides in the debate and was praised by members on both sides for the quality of its work. The panel's recommendations are the culmination of a year-long investigation of the care of laboratory animals in the city (see *Nature* 338, 534; 13 April 1989).

The proposal that the city should appoint a Commissioner of Laboratory Animals is the most dramatic of those incorporated in the new ordinance. The commissioner will have authority to "make unannounced visits to inspect animal and research facilities as needed". The commissioner will make at least one annual visit to each research institution in the city, according to the recommendations.

The proposed ordinance will also require all research institutions in the city to conform to federal statutes and regulations concerning the care of laboratory animals. At present, most of the city's laboratories merely conform voluntarily to federal guidelines. The ordinance will also require that each research institution maintains an "autonomous animal care and use committee with the power to disapprove or restrict research" in accordance with federal laws. Each committee will have to include a member who is not affiliated with the research institution.

John Moses, an MIT physician who heads the university's animal care committee, and who represented the scientific

community on the advisory panel, calls the proposed ordinance "a necessary step to give the city and the local community the accountability that they need" on the treatment of laboratory animals. He says the panel's recommendations are "very simple", "inexpensive" and "easy to live with", adding that they could help build public trust and stave off more rigid community regulation. "If researchers have to spend a few extra hours a year to achieve greater public confidence", he says, "it's well worth it."

Moses' sentiments were echoed by several other researchers in the area, including James Fox, Head of MIT's Division of Comparative Medicine. Fox says that "the biomedical research community is certainly amenable" to peer review of animal care. "If it makes the Cambridge community more comfortable", he adds, "I think that it is acceptable."

Although the thrust of the new ordinance reflects the panel's recommendations, many key details are still to be decided. Differences remain on the method by which the city commissioner will be selected and on the question of



whether animal care committees should be required to include an animal welfare advocate. These details are expected to be decided early this summer when the City Council votes officially on the new regulations.

Cambridge Mayor Al Vellucci, who gave his public endorsement to the panel's recommendations and announced before last week's crowd that he will vote for the ordinance when it is written, likened the current situation to the recombinant DNA debate in Cambridge ten years ago in which he played a key role. Vellucci told the packed meeting that in the former debate, "Nobel laureate scientists told us how damaging our regulations would be to their research; but once we passed it, they acknowledged that they could live with it". So, says Vellucci, will it be with the care of laboratory animals.

Seth Shulman

Just another meeting?

Montreal

LAST week's Fifth International Conference on AIDS ended without significant deepening of understanding of the pathogenesis of the disease or progress in developing a vaccine, and with a widespread feeling that for biomedical researchers the annual conference has turned into just another scientific meeting.

Public policy pronouncements and discussions of the sociological and behavioural aspects of the AIDS pandemic were more prominent than presentations of new biomedical research, and the focus was on how society will cope with the epidemic, given that biological science offers no immediate answers.

As at last year's AIDS conference in Stockholm, most of the biomedical presentations were updates of continuing research, with very few significantly new reports. Many epidemiologists seemed to feel that the heady days of rapid strides in knowledge have slowed to a more routine pace.

But many were less stoical, complaining that sessions dealing with similar topics had been scheduled simultaneously, and that the conference's nearly 12,000 participants and wide range of topics made meaningful exchanges between people in the same field impossible.

Robert Gallo, from the US National Cancer Institute, said that "we didn't expect this much diversity", adding that "we can't even find one another to talk."

Controversy surrounded a plenary address from the health commissioner of New York City, Stephen C. Joseph, who said that advances in the treatment of HIV-infected people that prevent their progression to disease may make it necessary to register infected individuals and to trace their sexual partners. But the new US Assistant-Secretary for Health, James O. Mason, said that it is important to "avoid contention" and that, until HIV-infected individuals are "comfortable" with the idea of contact tracing, it is more important to provide anonymous testing to avoid driving people away from diagnosis and treatment.

As the conference closed, there was some talk of dividing next year's AIDS conference — to take place in San Francisco — into separate meetings on biomedical advances and socioeconomic implications so as to make the programme more manageable. The Sixth International Conference on AIDS is expected to draw roughly 18,000 participants. Asked whether many scientists will avoid next year's meeting because the event has grown too big, Gallo hedged with "some will and some won't".

Carol Ezzell

Fasella offers to take advice

London

THE key words to describe the pattern of support for collaborative research in Europe after 1992 are "decentralized" and "flexible", according to Paolo Fasella, director-general of DG XII, the European Commission's directorate-general responsible for science policy. There will be no more "white elephants" in the form of new supranational European research centres; instead existing national laboratories will take on a "European function", perhaps for a limited period of time.

In a discussion meeting last week on the future of biological research attended by a select group of Britain's biologists and science policy-makers, Fasella rejected the idea of laboratories with a permanent "EC label". Research would be more efficient and effective if research institutes currently recognized as national centres of excellence were awarded grants for three to five years, he said, during which time they would be expected to recruit the best of Europe's researchers in that field and cooperate with a network of "satellite" laboratories carrying out relevant research.

Fasella also outlined the Commission's ideas on an "Assembly of Scientists and Technologists" to act as an independent advisory body to the Commission (see *Nature* 338, 734; 27 April 1989). Fasella said the Assembly would consist of 100–200 advisors loosely grouped together to form a self-governing group.

Although members would receive no salary, just "tourist fare travel expenses and sandwiches", says Fasella, they would be expected to initiate enquiries as well as to respond to requests for advice, and the Assembly would be funded by DG XII's current budget for paying consultants. This parsimony was strongly criticized by many at the discussion meeting. Expressing a common view, Professor John Ziman of the Science Policy Support Group said that the Assembly could not be successful without adequate finance and administrative support. But Fasella's broad plan to improve the "haphazard" manner in which advice on scientific and technological matters is provided at present was warmly welcomed.

The trade war with the United States after a European ban on steroid hormones in meat was cited as an example of the problems that can result from inadequate advice to European policy-makers. And unnecessarily tight regulations on the levels of nitrates in water "had a good chance of wiping out Europe's horticultural industry" because of high costs, complained Professor Bill Stewart, head of Britain's Agricultural and Food Research Council.

Fears about manpower haunt many discussions among British researchers about

the future of the biological sciences in Britain, and last week's discussion was no different. Departmental heads say that for the first time it is difficult to fill research studentships.

In an attempt to draw attention to this problem, a group of biologists led by Professor C.F. Higgins of the University of Dundee, sought endorsement from academics and industrialists for a letter to the Secretary of State for Education and Science, Kenneth Baker, and the heads of the research councils, in which they urge for action to be taken to improve the salary and career structure for young researchers. The poor prospects for researchers is causing graduates to choose other professions and creating an "internal brain-drain", the letter says. Signed by more than 100 heads of biological sciences departments, almost 700 tenured academics and 21 industrialists, it calls for "immediate action at the highest level" to find a "radical solution" to the problem.

Christine McGourty

Staab floats three ideas

Wiesbaden

MAX Planck Society President Heinz Staab put forward three proposals to strengthen scientific cooperation in the European Community (EEC) at the society's annual meeting here last week.

First, Staab proposed the institution of European research conferences modelled on the Gordon Conferences in the United States. Second, he recommended a European exchange programme for postdoctoral candidates to allow qualified people from one member country to carry out research at institutes elsewhere in the community. And third, on behalf of the German Research Ministry (BMFT), he announced a readiness to support collaborations between individual West German researchers and individual researchers in other countries.

The first two programmes would be supported by the European Community, if funds are available, but administered by the European Science Foundation. The conferences would be held "away from the city", he said, just like the Gordon Conferences in New Hampshire. Each conference would bring together about 100 persons in a narrow subject area for lectures and discussions.

Staab reported that both DG XII's Paolo Fasella and Maria Filippo Pandolfi, Vice-President of the Commission for Science, Research and Development, were receptive to the idea.

Steven Dickman

Ariane gets heavy

AFTER a 10-day delay, the European commercial satellite consortium, Arianespace, launched its largest Ariane-4 rocket so far on 5 June. On board were two telecommunications satellites, the US-built Superbird-A for the Japanese Space Communications Corporation and DFS Kopernikus-1, belonging to the West German Deutsche Bundespost. The Ariane-44L needed a total lift performance of 4,418 kg, putting it in the 'heavyweight' class that Arianespace needs to keep its commercial edge in the 1990s, in the face of a return of US launchers in the private sector. Chinese and Japanese launchers will also be seeking a slice of the \$1,000 million-a-year commercial launch market.

This year, the Ariane series celebrates its tenth anniversary. With 13 successful launches in the past 21 months, Arianespace has shown that it has recovered from the technical difficulties of 1986 which held up launches for over a year.

P.C.

AIDS white paper

THE Australian Government released its first white paper on AIDS last week, recommending that all foreigners wishing to settle in Australia be tested for AIDS and that the possession of small quantities of illegal drugs for personal use be decriminalized. The policy document will be presented to the Cabinet after a period of public consultation lasting a year. Other recommendations include establishing needle-exchange and disposal centres in all states, testing all prisoners for the AIDS virus, making condoms available in prisons and banning home AIDS-test kits.

Its most controversial recommendation is for a review of Commonwealth censorship and broadcasting laws. The review is designed to prevent a repeat of an incident earlier this year when announcements giving information about AIDS had to be altered substantially because television stations deemed them offensive.

T.E.

Taking the biscuit

KOALA bears may be popular at zoos throughout the world but they are not easy to keep, given that all they will eat are the leaves of a tiny number of eucalyptus species. But air delivery of fresh Australian eucalyptus leaves to the world's zoos may soon be at an end, thanks to the development of the koala biscuit by Professor Ian Hume of the University of New South Wales and Dr Lester Pahl of the University of New England.

The biscuit is similar in physical and chemical characteristics to the fresh leaves of the three most preferred eucalyptus species. The only problem is the koalas, which Hume accuses of being "fixated with their gum leaves". "Educating the koalas to eat the biscuits is a very long and painful process", he says.

T.E.

National identities blur

Tokyo

THE United States is still forging ahead of Japan in the development of biotechnology, according to two recent US reports. But Japan is the United States' leading competitor, particularly in the application of biotechnology to the development of new drugs. Such comparisons of the two countries, however, are beginning to lose meaning as more and more US and foreign companies set up research and development operations in Japan.

A survey by the US Federal Drug Administration (FDA) of Japanese companies that use biotechnology to produce pharmaceutical products found 9 products approved for marketing and 143 in various stages of research and development. The most intensive efforts are concentrated on agents that prevent or break up clots; vaccines to prevent various viral infections; interleukins to stimulate patients' immune response; colony stimulating factors to enhance the numbers of various cells present in the blood at lower than normal levels; and interferons for treatment of cancers or viral infection.

Cancer is a major focus of Japanese research with well over half of the products directed at cancers or related conditions. But, although Japanese production of new drugs using biotechnology is "robust", the United States holds roughly a two-to-one lead in both the number of companies involved and number of products, according to the FDA.

A survey by the US Pharmaceutical Manufacturers Association of biotechnology patents issued in the United States between 1986 and 1988 reaches a similar conclusion. Of the 237 patents issued for genetic-engineering techniques, US companies account for 195, Japanese companies for 22 and Western European 13.

Mitsuru Miyata, editor-in-chief of the Japanese newsletter *Nikkei-Biotechnology*, adds the tremendous effort that Japanese companies are now expending on the development of monoclonal antibodies to the FDA list. But he questions the value of comparative studies, given the growing multinational character of industrial research and development.

In the past, US companies licensed techniques in biotechnology to Japanese companies. But the recent trend is for US and foreign companies to carry out their own clinical trials and research and development in Japan. A good example is growth hormones where early products were developed by Japanese companies with licences from foreign companies; Eli Lilly and Serono Laboratories are now conducting their own clinical trials in Japan, according to Miyata.

The phenomenon is not confined to biotechnology. The huge chemical com-

panies DuPont and ICI recently set up research institutes in Japan, as has the West German drug manufacturer Hoechst and the Swiss giant Ciba-Geigy.

IBM has long had basic research facilities in Japan, and LSI Logic, a US semiconductor company, has built a 1,000 million Yen (\$7 million) technical centre at Tsukuba, north of Tokyo. And while foreign companies set up in Japan, Japan's industrial giants are rapidly building research institutes overseas (see *Nature* 338, 697; 27 April 1989), making national classifications a thing of the past.

David Swinbanks

■ A shift towards cooperative industrial research wins backing in a joint Japan-US report* issued last week by the US National Academy Press. The report has

its origin in a meeting of 40 experts organized last summer by the US National Science Foundation and the Japan Society for the Promotion of Science.

Two broad areas where cooperative efforts might pay off were identified: definition of a computer-aided product realization system to help streamline the process of turning innovation into product, and basic research in intelligent manufacturing control systems to provide the scientific foundations for a totally automated, 'intelligent' factory. A third area of inquiry, on how universities and industry coordinate their research efforts, was scheduled for further discussion. The committees will meet again next year to present and compare plans for specific joint research projects. **Alun Anderson**

**Manufacturing Research Exchange: Foundation of a Japan-US Cooperative Research Program. National Academy Press 1989.*

NAZI VICTIMS

Brain sections to be buried?

Frankfurt

COINCIDENT with its fortieth annual meeting, the Max Planck Society (MPS) last week announced plans to give a respectful burial to tissue samples from the brains of Nazi euthanasia victims that remain in its collections. According to Director Heinz Wässle of the Max Planck Institute for Brain Research (MPIBR) in Frankfurt, the brain tissue — fixed in thin sections on up to 10,000 glass slides — will be cremated and the ashes will be buried at an appropriate site, perhaps an existing Holocaust memorial.

The tissue samples were collected by neuropathologist Julius Hallervorden (1882–1965) from the euthanasia centre at Brandenburg-Görden. Hallervorden, a section leader at the Kaiser Wilhelm Institute for Brain Research in Berlin, the forerunner of MPIBR, received 697 brains from the centre between 1940 and 1944. Thirty of the tissue samples have been shown to derive from these brains; up to 400 samples have been linked circumstantially with the killing centres.

The Max Planck Society will cremate all slides and samples in its possession dating back to the years 1933 to 1945. The permission from University of Frankfurt is required, but is expected to be granted.

A scandal erupted in Israel in January when a West German television report revealed the presence of tissue samples and skeletons of Nazi victims at the West German universities of Tübingen and Heidelberg (see *Nature* 337, 195; 1989).

The University of Tübingen appointed a commission to investigate the presence of such samples at the university and the possibility that they had been used for teaching. The commission is to meet soon and is expected to consider burial. The

University of Heidelberg is "still deciding" how best to dispose of four samples in its anatomical collections that may have been taken from Nazi victims, said spokesman Michael Schwarz. The samples may be buried at a memorial to Holocaust victims in a Heidelberg city cemetery. Neither university had heard of the MPS plans.

Wässle says that researchers have not used the Hallervorden samples in recent years, and that he and his co-director, Wolf Singer, did not know of the collection's existence when they arrived at the institute eight years ago. They learned of it only when historian Götz Aly tried to gain access to it in 1983 and "locked it away" in 1987 when Aly's book offered proof of its origins.

Cremating the collection respectfully seems to be the "only ethical solution", said Wässle, but he is also concerned that the reputation of science might suffer from further publicity. Meanwhile, the slides are locked in the basement room T0011 and are accessible only with a key stored in the director's desk.

The brains used by Hallervorden were removed from the cadavers of children who allegedly suffered from psychiatric disorders. Psychiatrist and historian Robert Jay Lifton estimates in his book *The Nazi Doctors* that 5,000 children were killed in the official Nazi euthanasia programme. Hallervorden continued to publish papers about his findings until his retirement in the early 1960s.

Geneticist Benno Müller-Hill, who collaborated with Aly's investigation some years ago, said that it is important for the records of the collection to be retained even after the samples are destroyed. **Steven Dickman**

UK law in the offing

London

JUMPING the gun on a long-awaited report from the Royal Commission on Environmental Pollution, the UK Department of the Environment last week issued a consultation paper proposing new legislation concerning the deliberate release of genetically manipulated organisms. Should the proposals become law, releases would be subject to notification and consent and, on the now fashionable 'polluter pays' principle, the costs of the consent process and those of any environmental damage resulting from the release would be borne by the applicant.

The proposed legislation is intended to plug gaps in existing laws. A major gap exists for releases where damage to the environment is the paramount risk. In particular, The Wildlife and Countryside Act (1981) makes it an offence to release any new animal into the wild but it is doubtful that the term 'animal' can be stretched to encompass bacteria or viruses, and it is equally doubtful that the addition of a single gene creates a 'new' animal, despite the recent designation of a fruitfly with an added heat-sensitive gene as such.

New legislation would include provision for notification of planned experiments to the Department of Environment, an authorization process and enforcement measures. There is some indication, however, that exemptions might be granted for certain categories of release which could grow in number as experience was accumulated. An attempt to devise such categories based on hazard analysis is under way but is "not within shouting distance of success", according to one of its advocates, Professor John Beringer of Bristol University.

The consultation paper suggests that genetically modified organisms imported as finished products should be treated no differently from those that had been subject to consent in the process of national development, regardless of the regulatory hurdles they may have crossed in their country of origin. (European legislation for releases — see *Nature* 339, 413; 8 June 1989 — will have to consider this problem within the context of the European Community.)

Another suggestion is that a register of applications for consent is published in the public interest, despite strong arguments in favour of confidentiality, especially for commercial reasons.

Comments on the consultation paper are requested before the end of August to allow new legislation to be included in a general Environment Bill that may be ready later this year. It is suggested, however, that comments are delayed until after publication of the Royal Commission's report on "The Release of Genetic

ally Engineered Organisms to the Environment", whose overly long gestation period is expected to come to an end early in July. The commission's chairman, Lord Lewis, guardedly says of the new proposals only that "in broad terms, they are not inconsistent with the Commission's views on the sort of legislation that is required".

NUCLEAR POWER

Rancho Seco shutdown vote

Boston

A MAJORITY of the residents of Sacramento, California, voted last week to shut the 918 MW Rancho Seco nuclear reactor. This is a serious blow for the US nuclear industry. Similar referenda have previously been held in numerous states, but this is the first vote in favour of shutting down a working nuclear plant.

The 53.4 to 46.6 per cent vote turned more on economics than on worries about safety. The Rancho Seco plant, completed in 1974, has consistently operated well below average capacity for nuclear reactors in the United States and, in 1988, produced electricity at roughly twice the price per kilowatt hour of electricity from other sources.

Opponents of Rancho Seco have nevertheless hailed their victory as a "shot heard around the world", and have hastened to add that voters rejected the plant in spite of more than half a million dollars spent by the nuclear industry on its lobby to keep the reactor running.

The nuclear industry, on the other hand, prefers to paint the referendum as a "unique situation". Scott Peters, of the Council for Energy Awareness, the nuclear industry trade association, says the vote was "not so much against nuclear power as against a poorly operating plant".

Peters stresses that Rancho Seco was run by a small public utility whose board of directors has itself been divided over the continued operation of the plant. And he notes the irony that the utility has invested \$400 million within the past three years on a "mass refurbishing" of the reactor. The recent vote, he says, does nothing to address the problems of pollution that will be generated by nuclear power's alternatives, or to meet the increased electricity demand faced in several parts of the country.

Richard Rosen, of the Boston-based Energy Systems Research Group, predicts that the economic problems that topped Rancho Seco are not unique. Rosen says that smaller reactors in particular have difficulty generating electricity at competitive prices because of high opera-

Anticipating eventual legislation, the Department of the Environment already operates an interim expert committee to provide advice on environmental aspects of planned releases. The Health and Safety Executive's Advisory Committee on Genetic Manipulation also takes environmental hazards into account when offering advice on planned releases, whose notification to the executive is soon to be made compulsory.

Peter Newmark

ting and maintenance costs relative to their output. Peters and others dispute this contention. But with the announcement earlier this year that a Colorado utility would shut its poorly-operating Fort St Vrain reactor for economic reasons, and the abandonment of the Shoreham reactor in Long Island, New York, Rancho Seco's closure is another major setback for the nuclear industry.

What will become of the Rancho Seco reactor remains to be determined — but there is no permanent disposal site in the United States for high-level nuclear waste, including Rancho Seco's nuclear fuel.

Seth Shulman

Nuclear revival hopes

Washington

A TOTAL of 417 nuclear reactors produced 17 per cent of the world's electricity last year, according to the annual survey of the US Council for Energy Awareness (USCEA). France took the lead with 70 per cent of its electricity generated at nuclear power stations, followed by Belgium at 65 per cent. But in terms of the total number of power stations the United States remains the world leader with 111 plants producing 20 per cent of electricity generated. In the Soviet Union 90 plants produced 11 per cent of electricity.

A comparison with 1987 data shows that the number of nuclear plants under construction fell, largely because of cutbacks in

Nuclear programmes worldwide

	1988	1987
Operable	417	404
Under construction	114	126
On order	16	9
Planned	96	149

the Soviet Union. The number of planned nuclear plants also fell as Venezuela, the Soviet Union, Yugoslavia and Israel abandoned tentative projects.

In May, nuclear power operators formed the World Association of Nuclear Operators to help improve safety and pool information. US operators look to the greenhouse effect and concern over the burning of fossil fuels to restore 'nuclear's' fading popularity.

Alun Anderson

A spin-labelling Nobel?

SIR—*Perestroika* (restructuring) and *glasnost* (openness) in the Soviet Union have resulted in corresponding changes in the attitude of the world community towards the Soviet Union and the Soviet people. Does this process affect the Nobel prize committee? It is no secret that in the past there has been a prejudice against Russian and Soviet scientists as well as against scientific results coming out of Soviet laboratories. For example it is well-known that Dmitriy Mendeleev lost out by only one vote to the discoverer of fluorine, Henry Moissan, winner of the Nobel prize in 1907. Others passed over by the committee include Alexander Arbuzov, the founder of the chemistry of organophosphorus compounds (although Brown was awarded a Nobel prize for his studies of organic boron compounds); Alexander Braunstein, who elucidated the transamination reaction; Eugenyi Zavoisky, who discovered the phenomenon of electron para-magnetic resonance (a later work on nuclear magnetic resonance was marked by a Nobel prize). The Nobel prize committee honoured C. Raman instead of Mandelstam. The only Soviet Nobel prizewinner in chemistry was Nikolay Semenov, whose scientific discoveries were so significant that they could not be ignored.

My question is: can the Nobel prize committee now examine more carefully the discoveries made in the Soviet Union? The selection of prizewinners is up to the committee, but I would like to suggest the field of nitroxyl radicals and spin labelling as worthy of recognition. In this field, great results were achieved in the Soviet Union. Harden McConnell (Stanford University) was the first to work out spin labelling, and obtained important results in immunology and the study of membrane structure using this technique. However, the chemical basis of spin labelling is an unusual class of nitroxides, the nitroxyl radicals¹, and their nonradical reactions, discovered in the Soviet Union by Edward Rozantsev. (In France, André Rassat had also worked on nitroxyl chemistry.) Nitroxyl radicals have an 'extra' function, so can react without involving free valency, hence these novel free radical reactions². When McConnell worked out the theory of spin labelling, he used nitroxyl radicals and their nonradical reactions in the preparation of spin-labelled biomacromolecules³.

Nitroxyl radicals are also widely used in chemical physics, chemical kinetics, analytical chemistry, polymer chemistry, and even in pharmacology. A few thousand papers, many reviews, and monographs⁴⁻⁷ have been devoted to nitroxyl radicals and spin labelling. Has the Nobel prize committee taken note of this novel area of

chemistry (nitroxyls and spin labelling) and its founders: McConnell, Rozantsev and Rassat?

RENAT I. ZHDANOV*

*Institute for Biotechnology,
Moscow 117246, USSR*

*Present address: Max-Planck Institut für medizinische Forschung 6900 Heidelberg 1, Jean-Strasse 29, FRG

1. Neiman, M.B., Rozantsev, E.G. & Mamedova, Y.G. *Nature* **196**, 472–473 (1962).
2. Rozantsev, E.G. *Free Nitroxyl Radicals* (Plenum, New York, 1970).
3. Stone, T.J., Buckman, T., Nordio, R.L. & McConnell, H.M. *Proc. natn. Acad. Sci. U.S.A.* **54**, 1010–1014 (1965).
4. *Spin Labeling. Theory and Application* (ed. Berliner, L.J.) Vols 1–3 (Academic, New York, 1976, 1979, 1989).
5. *Spin Labeling Method. Problems and Outlooks* (eds Emanuel, N.M. and Zhdanov, R.I.) (Nauka, Moscow, 1986).
6. *Nitroxyl Radicals. Synthesis, Chemistry and Applications* (eds Rozantsev, E.G. & Zhdanov, R.I.) (Nauka, Moscow, 1987).
7. *Biomacromolecules in Spin Labeling* (eds Likhtenstein, G.I. & Zhdanov, R.I.) (Nauka, Moscow, 1988).

Hirohito's legacy

SIR—Emperor Hirohito's role during his 63-year reign will be debated for many years to come. The comment by Masakado Kawata and Eiich Kasuya (*Nature* **338**, 370; 1989) fails to address the basic issue.

I think that Hirohito really wished to remain a peaceful emperor biologist but did not have a courage to overcome the very strong imperial court system and the military power, which fed him only information advantageous to themselves. If the Emperor chose not to cooperate with the military, they could have fabricated excuses to justify replacing him with a younger brother.

But two questions remain: what lessons have the Japanese learned from the Pacific War, and how will the new Emperor Akihito behave in the years to come? Although the monarch does not have any political power under the new constitution, he is still very influential as the national symbol. In recent years, his vulnerability to being used by those on the fanatical right wing seems to have increased. After Japan's surrender in 1945, Hirohito declared that he was not a god, and travelled around the country meeting people. That practice has, however, long been discontinued except for very ceremonial occasions. Although Akihito had the courage and determination to marry a commoner and to send his son to the University of Oxford, he may not be able to break the ever-thickening 'chrysanthemum curtain'. If there is an attempt to use him for imperialist domination by military power, will he be able to resist? It remains to be seen how well-informed Emperor Akihito can keep himself and how fully he can express his thoughts.

SHIRO TANAKA

*11334 Pemmican Run,
Cincinnati, Ohio 45249, USA*

Raw data

SIR—Your report (*Nature* **339**, 163; 1989) on the investigation into a paper published in *Cell* raises serious concerns. The integrity of the scientific process requires that raw data be recorded fully, promptly and accurately and stored securely. There must be a clear distinction between recording of experimental details and observations, and jottings of ideas or sketches of proposed studies, which an individual may choose to keep in a more casual way (not advisable in relation to claims to priority). In any case of doubt about the relationship between the data and the published conclusions, whether fraud or error is suspected, there should be no question of not carrying out an audit of the raw data. To facilitate this, laboratory notebooks and data files should generally be the property of the institution, and the director of the institution should be responsible for ensuring their security.

Few would wish to see a system such as good laboratory practice for the pharmaceutical industry imposed upon the basic research community. But the integrity of scientific research must not only be maintained but must be seen to be maintained. As Ephraim Racker states (*Nature* **339**, 91–93; 1989), audits of data would have little effect on clever fraudsters, but such cases are thought to be very rare.

G. ALLEN

*26 Chesham Crescent,
London SE20 7RW, UK*

Is anybody there?

SIR—Speculation regarding the search for extraterrestrial intelligence (SETI), should be accorded a legitimate place in both science and *Nature*. Astronomers have educed both theoretical and observational evidence of planetary systems associated with stars, and now more than ever, intelligent life may be presumed by respectable (if hopeful) scientists to occur commonly in the Universe. Intelligent life, therefore, may indeed have evolved in planetary systems within striking distance of supernova 1987A. Although this supernova constituted a bonanza for Earth's astronomers, supernovae may tragically extinguish 'nearby' intelligent life.

The interstellar psychologists among us might suggest the search for extraterrestrial transmissions upon stars in the vicinity of stars approaching the supernova stage. Just as we have launched records of our existence into deep space, intelligent beings who may be aware of impending disaster might be more likely than others not so threatened to transmit messages to posterity.

ROBERT A. MICHAELS

*Ram Trac Corporation,
931 Northumberland Drive,
Schenectady, New York 12309, USA*

DNA fingerprinting on trial

Eric S. Lander

In the rush to use the tremendous power of DNA fingerprinting as a forensic tool, the need for standards has been overlooked.

WITH the exception of identical twins, no human beings have identical DNA sequences. Of the 3,000 million nucleotides which we inherit from each parent, about 1 in 1,000 is a site of variation, or polymorphism, in the population. These DNA polymorphisms are most conveniently detected when they alter the length of the DNA fragments produced by the action of restriction enzymes, giving rise to restriction fragment length polymorphisms (RFLPs). In standard practice, the length of the fragments is measured by the rate at which they move in an electrophoresis gel.

More than 3,000 RFLPs have been identified to date, including some 100 highly polymorphic loci at which dozens of variant alleles are present in the population. By using RFLPs to trace the inheritance of chromosomal regions in families afflicted with genetic disorders, human geneticists have been able to pinpoint the location of the genes causing diseases such as Huntington's disease, cystic fibrosis and others — in the process spawning the field of DNA diagnostics.

Forensic science has more recently latched onto RFLPs, but with a different purpose: to identify the individual origin of blood or semen samples found in criminal investigations based on their distinctive RFLP patterns. In the United States, forensic RFLP testing has been pioneered by two private laboratories — Lifecodes Corporation of Valhalla, New York and Cellmark Diagnostics of Germantown, Maryland — and is also used by the Federal Bureau of Investigation, which began testing earlier this year.

Since the first use of DNA 'fingerprinting' in a trial in Florida in 1988, DNA 'fingerprint' evidence has already been used in more than 80 criminal trials in the United States. Applying the legal standard for the admissibility of novel scientific evidence defined in *United States v. Frye* in 1923, trial judges have raced to admit DNA fingerprinting as evidence on the grounds that the methods are "generally accepted in the scientific community", citing the application of RFLPs in DNA diagnostics and accepting claims that false positives are virtually impossible.

With due respect, the courts have been too hasty. Although DNA fingerprinting clearly offers tremendous potential as a forensic tool, the rush to court has obscured two critical points: first, DNA fingerprinting is far more technically demanding

than DNA diagnostics; and second, the scientific community has not yet agreed on standards that ensure the reliability of the evidence.

DNA diagnostics requires simply identifying whether each parent has passed to a child the RFLP pattern inherited from his or her mother or father. Because the four discrete patterns are known in advance, these investigations have built-in consistency checks which guard against many errors and artefacts.

DNA fingerprinting, by contrast, is more like analytical biochemistry: one must determine whether two completely unknown samples are identical. Because hypervariable RFLP loci often involve 50–100 alleles yielding restriction fragments of very similar lengths, reliably recognizing a match is technically demanding. At one commonly used locus, for example, most alleles lie within a mere 2 per cent of the length of the gel.

Few molecular geneticists, in such circumstances, would declare a match without performing a mixing experiment, in which a 50:50 mixture of the two samples is shown to yield the same pattern as each sample separately. In the rare event that a mixing experiment could not be carried out, most molecular geneticists would at least insist on using internal controls — probes which detect non-polymorphic DNA fragments within each lane — to verify that the lanes have run at equal speeds and to provide standards against which fragment sizes can be measured precisely.

Yet the DNA fingerprinting results now being introduced into the US criminal courts are often based on much flimsier evidence. Not only are mixing experiments and internal controls often omitted, but some laboratories use no objective standards whatsoever for declaring a match.

Unlike DNA diagnostics, DNA fingerprinting also depends on inferences about the frequency with which matching RFLP patterns will be found by chance, which in turn rest on simplifying assumptions about population genetics whose accuracy has not yet been rigorously tested for highly polymorphic RFLP loci. For example, it is assumed without convincing proof that Caucasians, Blacks and Hispanics can each be regarded as homogeneously mixed populations, without significant subgroups, even when considering loci at which most alleles are relatively young from the perspectives of population genetics.

Yet despite such fundamental uncertainties, forensic laboratories blithely cite breathtaking frequencies: a recent report based on the study of only four RFLPs announced that the chance of an alleged match occurring at random was 1 in 738,000,000,000,000.

It is my belief that we, the scientific community, have failed to set rigorous standards to which courts, attorneys and forensic-testing laboratories can look for guidance — with the result that some of the conclusions presented to courts are quite unreliable.

My concern is not merely academic: during the past five months, I have been an advisor for the defence and given six days of testimony in what has turned out to be the longest and most searching pre-trial Frye hearing in the United States on the admissibility of DNA evidence — a murder case in the Bronx, one of the boroughs of New York City. I have also had occasion to investigate several other DNA fingerprinting cases in the course of preparing a report at the request of the US Office of Technology Assessment, to be delivered later this summer. (My own field is medical genetics: both projects arose as unintended consequences of accepting an invitation to a conference on DNA forensics at Cold Spring Harbor's Banbury Center in late 1988.)

It is my contention that DNA forensics sorely lacks adequate guidelines for the interpretation of results — both in molecular biology and in population genetics. To illustrate this, I will draw examples from cases with which I am personally familiar. I should emphasize that the focus on specific cases is not intended to criticize particular testing laboratories: with scientific consensus lacking, similar disagreements about interpretation would surely arise in many other cases involving other laboratories.

The Castro case

On 5 February 1987, Vilma Ponce and her 2-year-old daughter were stabbed to death in their Bronx apartment. Acting on a tip, police interrogated a neighbourhood handyman, Jose Castro. Detectives noticed a small bloodstain on Castro's watch, which was sent for analysis to Lifecodes. Company scientists extracted about 0.5 µg of DNA from the bloodstain, which they compared with DNA from the two victims.

The DNA was digested with the restric-

tion enzyme *Pst*I, size-fractionated on an agarose gel, and transferred onto a Southern blot. The blot was then hybridized with probes for three RFLP loci; DXYS14, D2S44 and D17S79, as well as a probe for a Y-chromosome locus to identify sex. (Human loci detected by random DNA probes are named according to chromosome and order of discovery.) On 22 July 1987, Lifecodes issued a formal report¹ to the district attorney (Table 1) stating that the DNA patterns on the watch and the mother matched, and reporting the frequency of the pattern to be about 1 in 100,000,000 in the Hispanic population. The report indicated no difficulties or ambiguities. Yet there are several fundamental difficulties, as follows:

DXYS14: Identifying bands. Contrary to the forensic report (Table 1), the only autoradiogram involving DXYS14 shows

who had provided the probe to Lifecodes, testified that the DXYS14 autoradiogram had to be considered to exclude the defendant in the absence of any experiments to explain away the non-matching bands.

Lifecodes' discounting of the two non-matching bands in the watch lane suggests that its identification of bands may have been influenced by making direct comparisons between lanes containing different DNA samples, rather than by considering each lane in its own right. Additional support for this hypothesis is provided by two further examples:

(1) In the dead daughter's pattern at DXYS14, all other expert witnesses for the prosecution and the defence identified only a single band (Fig. 1). However, Lifecodes' laboratory records show that it recorded three bands in this lane — in precisely the same positions as those recorded for the mother and the watch.

TABLE 1 Reported fragment sizes (in kilobases) at three RFLPs in *New York v. Castro*, as given in Lifecodes' formal report to the district attorney

	D2S44	D17S79	DXYS14
Blood from watch	10.25	3.87 3.50	4.83 3.00 1.94
Deceased mother	10.25	3.87 3.50	4.83 3.00 1.94
Deceased daughter	ND	3.87 3.50	4.83 — 1.94

five bands in the watch lane and only three bands in the mother's (Fig. 1). In his testimony, Michael Baird, Lifecodes' director of paternity and forensics, agreed that the watch lane showed two additional non-matching bands, but he asserted that these bands could be discounted as being contaminants "of a non-human origin that we have not been able to identify".

In my opinion, it is impossible to know whether the bands are non-human without demonstrating that they are absent when the experiment is repeated with an uncontaminated probe. How then did Lifecodes reach its judgment? Baird stated that, in his experience, DXYS14 should exhibit a pattern of fragments whose intensities decrease in proportion to their length: the extra bands could be ignored because their intensities were "not in the proportion I would expect to see".

In fact, the published scientific literature shows that DXYS14 actually yields patterns (ranging from one to more than six polymorphic fragments) whose intensities obey no ironclad rule. Ironically, the prosecution itself had put into evidence the very article which proved the point: the original paper² defining DXYS14, by Howard Cooke of the Medical Research Council in Edinburgh, containing photographs of *Pst*I Southern blots showing lanes in which hybridization intensity and fragment length are clearly uncorrelated (Figs 6c and 7a in ref. 2).

David Page of the Whitehead Institute later introduced *Pst*I blots from his laboratory showing arbitrary patterns of intensities at DXYS14. Finally, Cooke himself,

(For reasons I do not know, the forensic report listed only two of these bands.)

(2) Although D17S79 is expected to yield at most two bands, the dead daughter's lane exhibited four bands in the appropriate size range (not shown) in the only hybridization with the probe completed before the forensic report was issued. The report listed only two of these four bands — the two in the same position as in the mother and the watch.

The tendency to use lane-to-lane comparison to distinguish between bands and artefacts is perfectly natural; such comparison can be quite helpful in certain experiments. However, in my opinion, it is inappropriate in DNA fingerprinting analysis of unknown samples — as one runs the risk of discounting precisely those differences that would exonerate an innocent defendant. Forensic laboratories should be required to use objective criteria for identifying the bands in each lane, and to use experiments to rule out proposed artefacts.

When a result is reported to have an error rate of 1 in 100,000,000, it seems essential that the underlying data are not

left as a matter of subjective opinion.

D2S44 and D17S79: declaring a match. To obtain objective measurements of a band's position, Lifecodes uses a computer-digitizing apparatus^{3,5}. The approach is reported to be highly accurate: when identical fragments are electrophoresed in different lanes, the difference between their positions is reported to show a standard deviation (s.d.) equal to 0.6 per cent of molecular weight^{3,5}. Based on these experiments, Lifecodes defined a formal matching rule: two fragments are said to match when their positions differ by less than 3 s.d.s. The matching rule was explicitly stated in a recent population study⁵ ("two DNA fragments were considered to be of different size if their values differed by more than 3 s.d.s") and in the formal forensic reports¹ in the Castro case ("fragments with measurements that are within [3 s.d.s.] of each other . . . are considered indistinguishable and their average size reported".)

Because the fragments at D2S44 and D17S79 did not appear to match perfectly, the defence examined Lifecodes' computer measurements. In fact, the bands fell outside the declared matching rule; as shown in Table 2, the bands at D2S44 differ by 3.06 s.d.s and the lower bands at D17S79 differ by 3.66 s.d.s. Under the objective matching rule, the bands were non-matches.

Why then was a match declared? Lifecodes stated that it did not actually use the objective threshold of 3 s.d.s. for declaring a forensic match: its decisions were based on subjective visual comparison. Agreeing that the explicit statements in the forensic report implied that the objective criterion had been used, Baird allowed that the statement "may not be the best explanation" of the company's actual procedures. As far as I can see, there is also no mention of the use of visual matching in the company's scientific papers or forensic reports. Clearly, there has been a significant misunderstanding about the matching rule which Lifecodes has been using.

In my opinion, visual matching is inappropriate in DNA fingerprinting, inasmuch as (1) many alleles have very similar sizes; (2) the accuracy of the measurement process is reported to be known; and (3) without an objective definition of a match, there is no meaningful

TABLE 2 Measured fragment sizes (in base pairs) at three RFLPs in *New York v. Castro*, as shown in Lifecodes' records from its computer-digitizing apparatus produced in response to subpoena

	D2S44	D17S79	DXYS14
Blood from watch	10,350	3,877 3,541	4,858 2,995 1,957
Deceased mother	10,162	3,869 3,464	4,855 2,999 1,946
Difference (per cent of average size)	1.83	0.21 2.20	0.06 0.13 0.56
Difference in number of s.d.s	3.06*	0.34 3.66*	0.10 0.22 0.94

According to Lifecodes' published papers^{3,5} and its formal reports to district attorney¹, the difference between two identical bands in different lanes shows an s.d. equal to 0.6 per cent of fragment size. Asterisks, bands differing by > 3 s.d.s.

way to determine the probability that a declared match might have arisen by chance (see below).

DYZ1: Use of controls. The DYZ1 locus provides a convenient method of identifying the sex of a sample: the locus is repeated about 2,000 times on the distal long arm of the Y chromosome, giving rise to an intense 3.7-kilobase (kb) *Pst*I band in males and no such band in females. Based on a hybridization with a probe for DYZ1, the blood on the watch was said to have come from a female.

Indeed, the mother, daughter and watch DNAs showed no male-specific band. But neither did the lane marked control. Who was the control?

(1) Initially, Baird testified that the control DNA came from the female-derived HeLa cell line.

(2) Two weeks later, however, the technician who actually performed the experiments testified that the control DNA came not from HeLa cells but rather, he recalled, from a male scientist.

(3) Baird then explained the absence of a positive signal in the now-male control lane by telling the court that the male scientist has a "short" Y chromosome which "does not react with this repeat sequence", a condition which is "fairly rare, but it does happen". In conventional genetic terminology, the individual was a genetic mutant deleted for the region.

(4) I then testified that the population frequency of such complete deletions is about 1 in 10^3 – 10^4 , with most causing phenotypic abnormalities usually including complete sterility (D. Page, H. Cooke and K. Smith, personal communication). A normal male with such a deletion would be so rare as to be publishable.

(5) Baird then reported that the control DNA came not from the male scientist, but from a female technician. Although no precise record had been kept of which DNA preparation had been used, he said that he had managed to identify the source of the control lane by studying its RFLP pattern — an unforeseen use of DNA fingerprinting.

The confusion had probably resulted from faulty recollections (by Baird and the technician) and faulty inferences (about the male scientist), but it underscored the need for meticulous record-keeping in DNA forensics, which may not originally have been as clear.

Leaving aside the identity of the control DNA, there is a more important question: should a sex test be considered reliable without seeing a positive control on the autoradiogram to prove that the experiment had worked correctly? Baird testified that such a result could be considered "reliable". I would vigorously disagree.

D2S44: Analysis of degraded DNA. Based on seeing only a single band in the watch DNA, the sample was reported to be homozygous for a 10.25-kilobase (kb)



FIG. 1 Hybridization of probe 29C1 for the locus DXYS14 to *Pst*I-digested DNA from deceased mother (M), blood speck from defendant's watch (W) and deceased daughter (D), performed by Lifecodes in *New York v. Castro*. Although the autoradiogram shows three bands in lane M, five bands in lane W and one band in lane D, Lifecodes recorded three bands in identical positions in all three lanes. The autoradiogram is an overnight exposure; no further exposures or hybridizations involving this probe were performed.

band at the D2S44 locus. Although the conclusion may seem reasonable at face value, there is a serious problem: the small quantity of DNA on the watch was clearly degraded, and nearly 90 per cent of alleles in the Hispanic population lie above 10.25 kb.

How can one be sure that the sample was not a heterozygote, with a higher band undetected due to degradation? Estimating the extent of degradation by eye, Baird stated that the photographs of the ethidium bromide stain of the gel "gives you some indication that there is enough material present to be able to get a signal" in the 12–15-kb range", but that "I[d] hate to bet the ranch" on it.

Wouldn't a probe detecting a non-polymorphic single-copy band at about 15 kb have provided a definitive positive

control? He replied as follows:

"If you are making a decision based only on a single locus, whether or not a pattern you saw was homozygous or whether or not you are missing a band, you'd have to have some way to absolutely be sure that you were seeing everything that was there and the control that you mentioned would be very helpful to do that.

Now, in addition to the D2S44 locus, we also looked at two additional loci, D17S79 and DXYS14. By looking at the combination of loci, and seeing whether or not there was a pattern that matched or not, allowed us to determine or help[ed] us to determine that the pattern that we saw with the D2S44 locus is [a] homozygote."

Personally, I do not understand how the presence of matches at D17S79 and DXYS14 has any bearing on the determination of a match at D2S44: each test must be evaluated independently, especially as the individual probabilities of a match for each locus are multiplied together at the end (see below).

Probe contamination. To explain various artefacts, Lifecodes invoked four separate instances of probe contamination: human probes were said to be contaminated with bacterial sequences, and bacterial and plasmid probes were said to be contaminated with human sequences. Moreover, Baird testified that the company continued to use probes even after learning that they were contaminated, while apparently keeping no precise record of when such probes had been used.

Although the use of contaminated probes may be permissible in some types of experiment, it is, in my opinion, inappropriate in DNA forensics. Because the samples are of unknown origin and often contaminated, false matches can result but may be hard to recognize.

Population genetics. After declaring a forensic match, testing laboratories apply a three-step procedure to calculate the probability that the match might have arisen by chance in the population: (1) for each allele, one counts the frequency with which matching bands occur in a previously-drawn population sample (Lifecodes used a database⁷ of US Hispanics in the *Castro* case); (2) for each locus, one then computes the probability of observing a matching genotype by applying the classical Hardy–Weinberg equations^{5,6}, under the assumption that population is freely intermixing and thus contains no heterogeneous subgroups; and (3) for the complete RFLP pattern, one then multiplies the three single-locus probabilities, under the assumption that the genotypes at the loci are in linkage equilibrium (are uncorrelated). In fact, none of these steps stood up to scientific scrutiny in the *Castro* case.

Probability of a match: inconsistent matching rules. Whatever matching rule is used to declare a forensic match, it is axiomatic that the same matching rule must then be

used for counting the matches occurring in the population database. In fact, Lifecodes' calculations did not use the same matching rule.

■ To declare a forensic match with a given band, Lifecodes' published matching rule calls for examining a range of ± 3 s.d.s around the band. Obviously, the chance of a match arising at random is just the proportion of bands in this range.

■ To calculate the probability of a matching band occurring by chance, however, Lifecodes uses a completely different approach: the reported probability is essentially the frequency of bands occurring in a window of size $\pm \frac{2}{3}$ s.d.s. More exactly, as described in an as-yet-unpublished paper⁸, it is a weighted-average of frequencies corresponding to such intervals, with the terms weighted according to a gaussian distribution centred on the estimated allele size.

Because the method involves averaging frequencies corresponding to intervals that are 4.5-fold smaller than those allowed for declaring a forensic match, the probability reported for each allele will typically be too small by a factor of about 4.5. Because each RFLP involves two alleles, the probability may thus be understated by a factor of about $(4.5)^2$, or about 20, for each locus. For a three-locus genotype, the error may thus be about 8,000-fold. (Of course, these calculations assume that one uses the 3 s.d. matching threshold. If the less stringent standard of visual matching is used, the discrepancy may be even greater.) Such a statistical procedure is like catching a match with a 10-foot-wide butterfly net, but then attempting to prove the difficulty of the feat by showing how hard it is to catch matches with a 6-inch-wide butterfly net.

How does Lifecodes justify its approach? Once a forensic match is declared between two bands, Lifecodes apparently considers the average fragment size to represent the allele present in the sample. It then estimates the population frequency of this allele, essentially by counting the bands within a range so narrow that it may not even include either of the two actual measurements. This approach substantially underestimates the true chance of a forensic match occurring at random, as it takes no account of the actual threshold used for declaring matches.

Heterogeneity within the Hispanic population. To justify applying the classical formulas of population genetics in the Castro case, the Hispanic population must be in Hardy-Weinberg equilibrium. In fact, Lifecodes' own data show that it is not. The classical test for Hardy-Weinberg equilibrium is based on Wahlund's principle^{7,9} that the rate of homozygosity in a population containing distinct subgroups will be higher than would be expected under the assumption of random

mating. Applying this test to the Hispanic sample, one finds spectacular deviations from Hardy-Weinberg equilibrium: 17 per cent observed homozygotes at D2S44 and 13 per cent observed homozygotes at D17S79 compared with only 4 per cent expected at each locus, indicating, perhaps not surprisingly, the presence of genetically distinct subgroups within the Hispanic sample. (The expected 4 per cent frequency of homozygotes is based on the empirical probability of randomly drawing two alleles from the population sample that are either identical or so close together as to be scored as a single band; the minimum size difference needed to discriminate between one versus two bands in Lifecodes' experiments was stated explicitly in testimony and in a paper⁸.)

Once a population is known to be heterogeneous, one also cannot assume linkage equilibrium even for loci on different chromosomes: if an individual possesses an allele common among Puerto Ricans at one locus, it is more likely that he will do so at a second locus as well.

In fact, Lifecodes' population study³ is scientifically valuable. From an evolutionary point of view, highly polymorphic loci contain many young alleles which may not be uniformly distributed within the Caucasian, Black or Hispanic populations. Studies of RFLP frequencies can reveal the detailed substructure of the human population, shedding light on migrations and trading patterns.

Such complexities, however, can undermine the use of simplified calculations in DNA forensics. Without the assumption of Hardy-Weinberg equilibrium and linkage equilibrium, there is no reliable way to convert allele frequencies into overall genotype frequencies: applying the classical equations can lead to spuriously low probabilities of a match. Possible solutions include empirical studies to identify ethnic subgroups that are in Hardy-Weinberg equilibrium and theoretical studies to derive appropriate correction factors for heterogeneity.

The experts' statement

After stepping down from the witness stand in late April, I attended the next day a Cold Spring Harbor conference on genome mapping co-organized by Richard Roberts, who had been the prosecution's lead witness when the Castro case had begun in mid-February. As Roberts explained, his testimony had been intended simply to provide the court with a primer on DNA analysis. Concerned about the issues that had come to light, Roberts conceived a novel plan: a joint scientific meeting of the experts who had testified for either side to review the evidence.

At the meeting, held on 11 May, the experts agreed upon a consensus statement declaring that "the DNA data in this

case are not scientifically reliable enough to support the assertion that the samples . . . do or do not match. If these data were submitted to a peer-reviewed journal in support of a conclusion, they would not be accepted. Further experimentation would be required". In particular, the statement cited the inappropriateness of (1) discounting the extra bands at DXYS14; (2) declaring a match between bands at D2S44 and D17S79 whose measured positions differed by more than Lifecodes' announced threshold of 3 s.d.s; and (3) using a less-strict matching rule when declaring forensic matches than when searching the population database. (Baird was unable to attend the experts' meeting because of a prior engagement.)

At first, the prosecution indicated that it would withdraw the DNA evidence based on the advice of its own scientific experts. Eventually, however, the district attorney decided to press ahead.

The hearing then resumed, with former prosecution witnesses testifying now for the defence. The prosecution's efforts to mount a rebuttal case fizzled:

■ To rebut the non-matching bands at D2S44 and D17S79, the prosecution stated that the use of a new measurement system now reportedly showed that the bands actually lay within 3 s.d.s. The judge ruled this evidence inadmissible, however, calling a last-minute switch of methodology "highly unscientific".

■ To rebut the problem with degradation above 10 kb, Lifecodes probed the Southern blot with the human *Alu* repeat sequence and determined that it showed hybridization up to the 23-kb molecular mass marker. In my opinion, the experiment itself was meaningless (because the ability to detect a sequence repeated 300,000 times in the genome has no bearing on the ability to detect single-copy sequences), but it was unnecessary to explain this to the court. Defence attorney Peter Neufeld, by now a veteran reader of autoradiograms, noticed that someone had accidentally misread the size markers: the *Alu* hybridization actually extended only to the 9.8-kb marker.

■ To rebut the population-genetic issues, the prosecution made eleventh-hour phone calls to various scientists, but they refused to testify for the state.

In the end, the prosecution's rebuttal case consisted only of the contention that the DYZ1 sex-test hybridization had actually worked correctly — based on the fact that Lifecodes had located another experiment done on the same day which showed a male-specific band in a rape suspect. (This would be an adequate, if unorthodox, control if it could be proved that both hybridizations had been carried out in the same plastic bag.)

The hearing concluded on 26 May, about 15 weeks after it began. Lawyers for both sides are now preparing briefs. Jus-

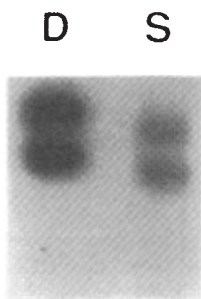


FIG. 2 Hybridization of probe pAC256 for the locus D17S79 to *Pst*I-digested DNA from defendant (D) and semen sample (S), performed by *Lifecodes* in *Georgia v. Caldwell*.

tice Gerald Sheindlin is expected to issue a decision in July. Sometime thereafter, the murder trial itself will commence — with or without the DNA evidence.

Other illustrative cases

The general issues of interpretation are unique neither to the Castro case nor to *Lifecodes*, as the following examples show.

Georgia v. Caldwell. In a death-penalty case currently in progress in Atlanta, James Caldwell stands accused of raping and killing his daughter Sarah. According to the forensic report, Caldwell's blood matches semen samples from the crime with 67,500,000,000:1 odds against the match arising at random. Testifying at the Frye hearing, *Lifecodes* scientist Kevin McElfresh described the process of declaring a match as a "very simple straightforward operation", asserting that "there are no objective standards about making a visual match. Either it matches or it doesn't. It's like if you walk into the parking lot and see two blue Fords parked next to each other. That's the situation here."

In fact, the patterns clearly do not match by eye (Fig. 2). McElfresh agreed, but asserted that "There is, however, a consistent non-alignment of the bands throughout the test, telling us there's a match." In other words, McElfresh contended that the differences were due to one lane having run faster than the other, although *Lifecodes* presented no internal controls to support this explanation.

Although the hearing had been expected to last only a week, prosecutors asked for a month-long delay to prepare their rebuttal. Observers speculated that *Lifecodes* might use the time to test an internal control. If so, it would raise the question of whether Frye hearings are becoming a substitute for having generally accepted laboratory protocols in the first place.

New York v. Neysmith. In a Bronx case in 1987, Hamilton Neysmith was charged with rape based on the victim's identification. Asserting his innocence, Neysmith hired *Lifecodes* to compare his blood with semen samples from the assailant: the laboratory declared an exclusion. Protest-

ing that the defendant may not have sent his own blood, prosecutors obtained a court order to compel a second blood sample: *Lifecodes* reported in August 1988 that the two blood samples came from different people. Based on this evidence, Bronx assistant district attorney Karen Yaremko asked the court to revoke Neysmith's bail, planning to charge him with obstruction of justice. The judge declined to revoke bail and the defendant, rather than leaving town, maintained his innocence, and demanded a third blood sample. After Yaremko pressed *Lifecodes*, she said, the company determined that an error had indeed occurred. Having come close to losing his liberty over inaccurate DNA results, Neysmith was finally exonerated after blood and semen samples were sent to Cellmark Diagnostics which confirmed the original exclusion. (*Lifecodes* declined last week to comment about the incident, which may have been nothing more than the sort of sample mix-up that can occur in any clinical laboratory. In view of the infallibility with which many jurors regard DNA fingerprinting, however, it may be that even stricter sample-handling procedures should be required.)

In *New York v. McNamara* in November 1988, another Bronx defendant sought to prove his innocence with DNA fingerprinting. Assistant district attorney Renee Myatt opposed the request, telling the court, "the office policy in dealing with a particular agency that does testing with respect to DNA [is] that their testing has been inaccurate, and therefore, unreliable." Notwithstanding this policy, the same District Attorney's office sought to introduce DNA evidence three months later in *New York v. Castro*.

Texas v. Hicks. In this rape-murder, the odds of the declared match occurring at random were reported to be 1 in 96,000,000. Apart from the issue of the matching rule used for searching the database, the population-genetic analysis took no account of the fact that the crime occurred in a small, inbred Texas town founded by a handful of families. The defendant was convicted and sentenced to death.

The case of the abandoned baby. When the president of an insurance agency in Ocean City, Maryland left her car to be towed for repairs to a garage in February 1988, the mechanic claimed that he had discovered a dead infant in the back seat. Although the woman insisted and a detailed medical examination confirmed that she had not been pregnant, Cellmark Diagnostics reported that its DNA analysis showed that she was the mother. (The evidence consisted of one hybridization with a mixture of four RFLP probes in which the woman shared four of eight bands with the child, as well as two hybridizations using probes detecting certain multi-locus

repeats.) Local papers reported the sensational news. No murder charges were filed, however, after the state medical examiner determined that the baby had been stillborn on about 4 February.

As it happens, the woman later gave birth to a full-term baby girl on October 24 — conceived on about January 29, according to sonograms carried out by her obstetrician. In view of the apparent contradiction, Cellmark last week invited a group of outside scientists to reanalyse remaining DNA samples from the baby.

Setting standards

Readers should not conclude from this article that DNA fingerprinting is not a powerful tool for forensic identification or that current testing labs are not competent: as in the early stages of any new technology, some difficulties are to be expected. Rather, there is an urgent need for the scientific community to agree on clear guidelines for the procedures and standards needed to ensure reliable DNA fingerprinting. Legislators should also consider whether licensing and proficiency testing should be required in forensics. At present, forensic science is virtually unregulated — with the paradoxical result that clinical laboratories must meet higher standards to be allowed to diagnose strep throat than forensic labs must meet to put a defendant on death row.

An appropriate start would be a US National Academy of Sciences committee, charged with preparing a report on guidelines for DNA fingerprinting. There is ample precedent: when voice-print evidence began to be introduced in the 1970s, the academy convened such a group to examine the technology. An academy study on DNA fingerprinting had been planned for last year, but was postponed indefinitely when the National Institute of Justice would not finance it. As one justice official told me, the study was unwelcome: scientists had done their part by discovering DNA; it was not their job to tell forensic labs how to use it. □

Eric S. Lander is in the Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA.

1. *Lifecodes' formal reports to District Attorney* (22 July 1987, 12 August 1988 and 6 February 1989).
2. Cooke, H. *Nature* **317**, 687–692 (1985).
3. Baird, M. et al. *Am. J. hum. Genet.* **39**, 489–501 (1986).
4. Baird, M. et al. *Adv. forens. Haemogenet.* **2**, 396–402 (1988).
5. Balasz, I., Baird, M., Clyne, M. & Meade, E. *Am. J. hum. Genet.* **44**, 182–190 (1989).
6. Cavalli-Sforza, L.L. & Bodmer, W.F. *The Genetics of Human Populations* (Sinauer, Sunderland, Mass., 1989).
7. Hartl, D.L. & Clark, A.G. *Principles of Population Genetics* (Sinauer, Sunderland, Massachusetts, 1989).
8. Morris, J.W., Sanda, A.I. & Glassberg, J.J. *J. forens. Sci.* (in the press).
9. Wahlund, S. *Hereditas* **11**, 65–106 (1928).

ACKNOWLEDGEMENTS. The Castro defence was a joint effort with my colleagues Howard Cooke, Lorraine Flaherty, Conrad Gilliam, Philip Green and David Page, together with attorneys Peter Neufeld and Barry Scheck. I also thank Richard Roberts for his efforts to help resolve this hearing and Carl Dobkin for participating in the experts' meeting. I am grateful to Susan Lee and Stephen Lincoln for technical assistance.

Making a university in the desert

New Mexico, which has an estimable university already, is hardly the state in which one would expect to find the embryo of a centre of excellence. But that is what a handful of brave people plan.

THE Sante Fe Institute has everything going for it except big bucks, but even the budget is beginning to look quite healthy. That seems to be the general opinion at what is admittedly high season for the institute, physically on the edge of what remains a sleepy tourist city. Contrary to general expectation, the town is not the Sante Fe Railroad's railhead, which is some miles away, while its function as the state capital of New Mexico is similarly disguised by the grand buildings left over from the old colonial rule (by Spain).

The plusses are easily appreciated. As in much else of the South-West of the United States, the sky is usually cloudless unless (this is one of the rainy seasons) it is delivering a torrent of warm water. There is a quaintness about the region; Indian villagers seem to keep up their traditional rain-dances as much for their own benefit as for that of onlookers, but also prudently offer their skill at drilling artesian wells as a service to incomers to the region. And every skyline is a surprise.

But the institute's most important geographical asset is the Los Alamos National Laboratory, thirty miles away on the top of its still-unspoiled mesa. It is not so much that there are substantial crumbs of financial comfort to be found beneath such a rich man's table (federal regulations see to that), but that people at Los Alamos seem wholeheartedly to have welcomed the plan to found a free-standing intellectual enterprise on their doorstep, partly because the South-West is, after all, a long way from elsewhere.

If George Cowan, the director of the institute, has his way, there should be ample opportunities for intellectual trading between the two establishments. The institute has been founded by people who are predominantly theoreticians — people such as Murray Gellman (Caltech) and David Pines (Illinois at Champagne-Urbana) are members of the Scientific Board. So is Cowan trying to make a replica of the theoretical institute at Santa Barbara? Not a bit of it. The trouble with Santa Barbara, he says, is that it's almost part of a university. Cowan, almost immodestly for such a personally modest man, says that the goal at Sante Fe is to make a university. Nothing less.

The model is the Rockefeller University in New York, which also began life as a mere institute — and which might still be so called had it not been for the exuberant immodesty of the late and visionary

Detlev W. Bronk. Cowan's dream is of a faculty of about 100 people, with two or three times as many colleagues — post-doctoral people and students. But he and his colleagues, theoreticians though they may be, are prepared to take an empirical view of the timescale — ten years from now or "when I'm no longer here" provide what might be called, in other contexts, an open reading-frame.

The implication is that, while the institute as it appears to have satisfied itself and its managers that it can keep on making progress, a little luck could make a big difference and a lot of luck could dramatically telescope the future. Students of academic entrepreneurship should perhaps visit before the beginnings of this enterprise have been buried by its success.

One thing that visitors will notice is that the institute has settled on an organizing principle that must command general attention: complex systems. Cynics may choose to score two cheap points at this stage. Complex systems, they may say, are defined by theoreticians as those that cannot now (or yet) be handled and, nevertheless, given that chaos and fractality have become fashionable, are all the rage.

Modest Cowan has a telling answer. Complex systems is not a synonym for "insoluble problems", but rather a pointer to a class of problems which are soluble with techniques that now exist. Their common ingredients are that one must begin with a kind of taxonomy, categorizing either the states of the system in interesting ways or the classes of soluble problems according to the techniques required. Then one must acknowledge that complex systems have a memory of a kind — their macroscopic future is determined not simply by their macroscopic present in what might be thought a Newtonian way, but by their microscopic present. Finally, and luckily, computation is coming to be the equal of all this complexity.

Last week's outcome of this optimism was the assembly of this year's summer school, which Cowan insists is misnamed. He intends that it should not be a school at which graduate students fill their notebooks with information to which their fellow-students will not have access, but rather an institution at which people will have to think for themselves; the last week of four (it seems a short time) is given over to a project or a problem.

Feigenbaum (of chaos) will be one of

those giving four formal lectures and staying on to answer questions. So will be Stephen Wolfram, the ex-Caltech ex-Princeton alumnus whose strong suit of cellular automata appears to satisfy all of Cowan's criteria for a complex system. Is it a disaster that these people will find themselves speaking in rented space in the parochial (= "Catholic") school across the road?

That, they should reflect, is what academic entrepreneurship is all about. Last year (1988), the Sante Fe institute spent about \$800,000. This year, it is planning to spend about \$1.5 million (but about \$250,000 remains to be found, although everybody is confident that it will be). Next year, the plan is to spend about \$2.5 million on intellectual activity, some of which is already promised or effectively guaranteed by the efforts of those who are now part-time members of the faculty.

Meanwhile, the institute has a problem with which it is confronted every day: there is no place to sit. The rented *adobe* building that half-accommodates the institute is, for practical purposes, simply a covered corridor with offices on either side. Claustrophobics may be happier in the permanently-rented accommodation in the parochial school across the road. That every office appears to have its own Sun workstation is probably as much a measure of where people come from as of the strapping institute's largesse.

Cowan's immediate goal is to find the funds with which to build a better building, somewhere permanent. That is where the luck will matter. He needs a benefactor who must be (by the US federal government's definition) different from the federal government.

Otherwise, there is a risk that the institute's affairs will be too much determined by the people who have chosen to support what it is already doing. Citibank, for example, appears to have seen the light by recognizing that economics is also a complex system, but can the institute succeed if nearly half its budget comes from such a source, and if its programme is skewed accordingly? The vertebrate immune system satisfies Cowan's criteria of complexity, but nobody is paying for that connection to be explored. The lessons, for students of academic entrepreneurship, is to go now and then in ten years, so see how it has come out. It will be tragic if the desert has taken over.

John Maddox

Ribozyme self-replication?

Thomas R. Cech

CONSIDERABLE enthusiasm has been expressed for the idea that ribozymes — RNA molecules with catalytic activity — provided the basis for nucleic-acid self-replication early in evolution¹⁻⁴. Experimental tests of the feasibility of RNA-catalysed RNA replication confirmed that such reactions are chemically possible⁵. At the same time, they revealed some major obstacles. In no case had the catalytic centre of a ribozyme been found to act on an external rather than an internal RNA template, and there was a restrictive base-sequence requirement that was thought to be inherent to the system. On page 519 of this issue⁶, Doudna and Szostak announce

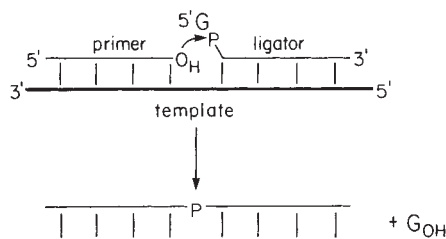


FIG. 1 Template-directed RNA ligation reaction catalysed by the *Tetrahymena* ribozyme⁶.

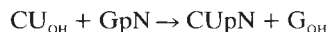
solutions to these two problems. These authors have engineered a ribozyme that stitches together oligonucleotide substrates to produce a strand complementary in sequence to that of an RNA template.

One of the defining characteristics of life is its ability to reproduce itself. At the molecular level, reproduction means the replication of genetic information. Contemporary cells accomplish replication through a complex interplay of proteins and nucleic acids. Considering the origin of life, one is immediately confronted with a dilemma: which came first, nucleic acid or protein? Information or function? The discovery that some RNA molecules can function as catalysts in addition to storing information has made catalytic RNA a plausible solution to this problem.

Although several distinct classes of ribozymes are now known, only the self-splicing introns have the nucleotidyl-transferase activity characteristic of contemporary RNA and DNA polymerases⁷. Many of the detailed models and experimental approaches to RNA-catalysed RNA replications have been based on group I introns and, in particular, on the ribosomal RNA intron of the protozoan *Tetrahymena thermophila*.

The RNA-joining activity of the *Tetrahymena* ribozyme has been known for some time. The reaction carried out by the ribozyme in nature, RNA self-splicing, involves the joining of a guanosine to the 5' end of the intron and the ligation of the two mature ribosomal RNA sequences, or

exons. When this ribozyme acts on other RNA molecules as substrates, it can perform various cleavage-ligation reactions. For example, Kay and Inoue⁸ found the *Tetrahymena* ribozyme to catalyse the following reaction:



The systems developed in these earlier studies shared a common feature: the nucleotide sequence that underwent ligation was determined by a template region or 'internal guide sequence' that was part of the ribozyme itself^{9,10}.

Doudna and Szostak⁶ use the ribozyme-catalysed reaction:



The reaction resembles that described by Kay and Inoue⁸, but the manner in which it is achieved is novel. The template region of the ribozyme has been removed from its catalytic centre, the portion that mediates the ligation reaction. The reactants align themselves on an external RNA template and are joined by the catalytic portion (Fig. 1), much as a modern DNA ligase joins two Okazaki fragments aligned on a DNA template.

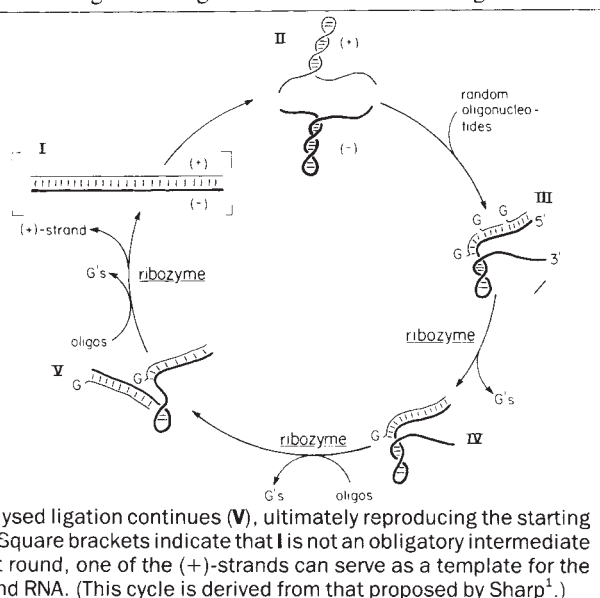
Considering this system in a bit more detail, in previous versions of the ribozyme, the internal template was tethered to the catalytic core by a covalent phosphodiester linkage^{5,7-9} or by base pairing of an adjacent sequence¹¹. In addition, it was presumably positioned and aligned by noncovalent tertiary structure interactions. (None of these has been identified, but the sort of interaction one imagines is hydrogen-bonding between the sugar-

phosphate backbone of the template and nucleotides within the core.) With the new version of the ribozyme designed by Doudna and Szostak⁶, the covalent linkage to the template has been severed, and there is no portion of the template that can base-pair to the ribozyme. The system must rely entirely on the noncovalent tertiary interactions to assemble the active complex. The authors conclude that these are relatively weak interactions, as evidenced by a high K_m (>0.1 mM) for binding of the template-substrate duplex to the ribozyme.

With the template separated from the catalytic centre, the ribozyme approaches being able to join any two oligonucleotides that line up by base pairing on the template strand, with no sequence restrictions. The one remaining impediment is the ribozyme's inherent proclivity for a uridine at the 3' end of the 'primer' oligonucleotide (see Fig. 1), and for a G in the template strand opposite the U. This U-G pair at the reaction site is highly conserved in group I introns, and it might have been an inviolable requirement for the ribozyme reaction. Reasoning that the importance of the U-G pair was the special geometry it provided at the reaction site, Doudna and Szostak tested a series of altered solution conditions to try to produce some subtle change in the RNA structure that might relieve the U-G requirement. The trivalent cation spermidine turned out to be the magic ingredient. In the presence of spermidine, ligation occurred with any Watson-Crick base pair preceding the ligation junction.

Having separated the template from the ribozyme and overcome the base-pairing constraint at the ligation site, Doudna and Szostak demonstrate the ability of their system to copy a template strand of arbitrary sequence. They find that the ribozyme can ligate as many as five oligonucleotides that are aligned on the

FIG. 2 Scheme for RNA self-replication, the fundamental steps of which have been demonstrated under controlled conditions in the laboratory. Double-stranded RNA (I), consisting of a ribozyme (+)-strand and a complementary (–)-strand, undergoes denaturation-renaturation. The (+)-strand folds to form the active ribozyme (II). Oligonucleotides (oligo), random in sequence except for their 5'-terminal G residue, assemble on the template (–)-strand by Watson-Crick base pairing (III). They are ligated together by the ribozyme (IV). Ribozyme-catalysed ligation continues (V), ultimately reproducing the starting RNA (I) and a new (+)-strand. Square brackets indicate that I is not an obligatory intermediate in the cycle². In a subsequent round, one of the (+)-strands can serve as a template for the synthesis of another (–)-strand RNA. (This cycle is derived from that proposed by Sharp¹.)



template by Watson-Crick base pairing. Products up to 42 nucleotides in length accumulate; only short reaction times are required, in the range of 5 minutes to 1 hour.

How would such a reaction fit into a complete scheme for RNA self-replication? One possibility is shown in Fig. 2. After one complete cycle of oligonucleotide ligation, a new ribozyme strand has been assembled. Assuming that the ribozyme could use a copy of itself as a template, a subsequent cycle of oligonucleotide ligation would accomplish the complete replication of the RNA.

Figure 2 also highlights some of the remaining obstacles to ribozyme self-replication. Ribozymes are of necessity highly structured RNAs, so a template for the ribozyme (minus strand in Fig. 2) would also be highly structured. To what extent would such structure inhibit replication? At what point would the guanine nucleosides released during ligation drive the reaction in the other direction, that of RNA cleavage? How would the very ends of the molecule be accurately replicated?

It is also necessary to consider the error rate of such a system. The 8–10 nucleotide primers and ligators used by Doudna and Szostak are long enough to pair stably with the template, but perhaps also long enough to tolerate internal mismatches between substrate and template. Furthermore, in the presence of spermidine, U-G pairs are still preferred over Watson-Crick pairs at the ligation site, and C-A and A-G pairs work well. Thus, the average error rate might exceed 20 per

cent per position. Although a prebiotic replication system would not be able to evolve if it did not mutate, neither could it be maintained if it were that promiscuous. If a ribozyme successfully established a self-replication cycle at the origin of life, perhaps it had a rather difficult active site that, like modern polymerases, used activated mononucleotides rather than oligonucleotides as substrates^{2,3}.

Finally, how would such a sophisticated ribozyme arise to get the whole process started? Presumably from something simpler. In this regard, it is encouraging that catalytic RNAs much smaller than the *Tetrahymena* ribozyme, containing of the order of 30 nucleotides per active unit, have been found to undergo specific cleavage reactions¹². And perhaps RNA itself arose from a simpler, nucleic acid-like molecule³. All of which goes to show that whenever one chicken-and-egg problem is solved, another might hatch to take its place. □

Thomas R. Cech is in the Howard Hughes Medical Institute, Department of Chemistry and Biochemistry, University of Colorado, Boulder, Colorado 80309-0215, USA.

1. Sharp, P.A. *Cell* **42**, 397–400 (1985).
2. Pace, N.R. & Marsh, T.L. *Orig. Life* **16**, 97–116 (1985).
3. Orgel, L.E. *J. theor. Biol.* **123**, 127–149 (1986).
4. Watson, J.D. *et al. Molecular Biology of the Gene* 4th edn, 1098–1163 (Benjamin/Cummings, Menlo Park, 1987).
5. Been, M.D. & Cech, T.R. *Science* **239**, 1412–1416 (1988).
6. Doudna, J.A. & Szostak, J.W. *Nature* **339**, 519–522 (1989).
7. Zaig, A.J. & Cech, T.R. *Science* **231**, 470–475 (1986).
8. Kay, P.S. & Inoue, T. *Nature* **327**, 343–346 (1987).
9. Been, M.D. & Cech, T.R. *Cell* **47**, 207–216 (1986).
10. Waring, R.B. *et al. Nature* **321**, 133–139 (1986).
11. Szostak, J.W. *Nature* **322**, 83–86 (1986).
12. Uhlenbeck, O.C. *Nature* **328**, 596–600 (1987).

PALAEOCLIMATOLOGY

Ice-age clues for warmer world

David A. Peel

THE latest computer-modelled predictions of greenhouse warming suggest that global mean air temperatures may rise by 5 °C over the next 30 years, with amplified rises of up to 12 °C in polar regions. This is comparable with the temperature increase from the last glacial period to the present interglacial, and the projected rate of increase is probably greater than at any time since then. Most global climate models have only tackled an equilibrium response to stepwise increases in atmospheric CO₂ concentrations, whereas in reality these increases are continuous. The results may be misleading both in terms of predicting the overall change and in determining regional climatic patterns. Evidence from the past, during periods of exceptional warmth or where temperatures were increasing, could help us to understand the characteristics of a future warm climate and will provide essential hard data needed to test whether com-

puter models can adequately simulate climate during periods of rapid change. New ice-core data presented by Dansgaard *et al.* on page 532 of this issue¹, from Dye 3 in southern Greenland, focus on an extremely rapid warming event which marked the end of the Younger Dryas oscillation 10,700 years BP (before present).

The Younger Dryas was the last in an apparently long series of climatic oscillations during the last glaciation, that were associated with 5–6 °C temperature shifts in Greenland². Rapid warming at the end of the event accounted for more than half of the overall glacial-interglacial warming. Evidence for it is found mainly in pollen records from bog and lake sediments in western Europe and from stable-isotope records in Greenland ice cores, although traces of the event are also recorded in eastern Canada. Therefore it seems probable that it was a phenomenon involving regions influenced by the North

Atlantic, as proposed by Mercer³, rather than a global event. Indeed, the final post-glacial warming seems to have been established in Antarctica well before this date⁴. It is possible, however, that a small 'cool' feature in Antarctic ice core records could be linked to the Younger Dryas.

Dansgaard *et al.*¹ present new details on the stable-isotope and dust composition of Greenland ice through the Younger Dryas termination, that sharply narrows the time span of this warming event to a mere two decades — equivalent to 1 metre of the Dye 3 ice core. Moreover, by comparing trends in the ¹⁸O and deuterium content of the ice, they can estimate the probable source temperature and hence oceanic latitude of origin of the moisture ultimately precipitated on the ice sheet. The authors suggest that the rapid warming was linked to a rapid retreat northwards of the polar front, with an associated retreat of the ice edge in the North Atlantic.

The rapidity of the event seems to lend further support to the idea that there was a bistable climate regime in the North Atlantic region during the glacial period, possibly connected to a flip-flop mechanism in the North Atlantic Ocean⁵. Broecker *et al.*⁶ have suggested that the Younger Dryas oscillation, together with the earlier glacial oscillations, could have been induced by a switching between 'strong' and 'weak' modes of deep water production in the North Atlantic. This process may have been driven in part by changes in the influx of cold, fresh water from the melting continental ice sheets. The rapid retreat of the ice edge now indicated by Dansgaard *et al.* provides a mechanism both for driving a significant influx of heat into high latitudes and for establishing enhanced evaporation rates in the North Atlantic — factors that could help to trigger the North Atlantic current into a strong mode.

Once the warming was initiated — and the external forcing factors involved still remain obscure — various feedback mechanisms could have amplified the effects of the northward displacement of the polar front. These include a decreasing meridional temperature gradient, reduced albedo in the formerly ice-covered areas of the North Atlantic, and an enhanced production rate of CO₂ in the North Atlantic owing to an increase in nutrient content of the polar surface waters. Indeed, there is evidence⁷ for sharp reductions in CO₂ concentrations during the Younger Dryas (and also in earlier oscillations) from both the Dye 3 and Camp Century ice cores. But recent analysis⁸ of the Byrd ice core from Antarctica has not revealed similar fluctuations. These conflicting findings are hard to reconcile with rapid global mixing of atmospheric gases and further confirmation is needed.

How relevant is the Younger Dryas termination event to the issue of greenhouse

warming? Clearly, many of the conditions and processes that were active 11,000 years BP differed substantially from the present. North America and Scandinavia were still extensively glaciated and the seasonal and spatial distribution of incoming solar radiation was radically different. However, if mode switches in the climate regime were important at that time then we must consider the possibility that they may also be involved in a future CO₂-induced super-interglacial. We must understand how changes in North Atlantic deep-water formation rates were implicated in these major climate shifts, as this process may play an important part in modifying, perhaps intensifying, greenhouse-gas-induced changes especially in the North Atlantic region.

Ice cores may be our only source of evidence available both to reconstruct vital details of rapid climate fluctuations in the past and to characterize changes in the ice-ocean-atmosphere system in the preceding periods, which may help to identify external forcing factors. For this reason a large international project is now underway to drill through the central part of the Greenland ice sheet. The drilling

site 'Summit' has been chosen because it should yield the longest (potentially more than 200,000 years) and most highly resolved palaeoclimate record possible in the Northern Hemisphere. Research teams from Europe and the United States start drilling this summer and plan to reach bedrock (about 3.1 km deep) in 1992. Cores will be drilled independently some 20 km apart but co-ordinated analysis will greatly improve the prospects for achieving unambiguous, accurately dated climatic signals. □

David A. Peel is at the British Antarctic Survey, Natural Environment Research Council, High Cross, Madingley Road, Cambridge CB3 0ET, UK.

1. Dansgaard, W., White, J.W.C. & Johnsen, S.J. *Nature* **339**, 532–534 (1989).
2. Dansgaard, W. & Oeschger, H. in *The Environmental Record in Glaciers and Ice Sheets* (eds Oeschger, H. & Langway, C.C.) *Dahlem Workshop* **8**, 287–317 (Wiley, Chichester, 1989).
3. Mercer, J.H. *Arctic Alp. Res.* **6**, 227–236 (1969).
4. Dansgaard, W. et al. *Geophys. Monogr.* **29**, 288–298 (1984).
5. Broecker, W.S., Peteet, D.M. & Rind, D. *Nature* **315**, 21–26 (1985).
6. Broecker, W.S. *Nature* **328**, 123–126 (1987).
7. Neftel, A., Oeschger, H., Staffelbach, T. & Stauffer, B. *Nature* **331**, 609–611 (1988).

CRYOBIOLOGY

Resurrection of frozen animals

Jared M. Diamond

SCIENCE FICTION enthusiasts dream of discovering a Neanderthal man frozen into an alpine glacier, thawing him, and observing how he regains life and behaves. This dream is not necessarily as silly as it sounds — many high-latitude invertebrates routinely survive freezing in the winter. The recent spectacular discovery of the first reptile to survive natural freezing¹, together with similar discoveries for four species of frog^{2–6}, at first raises hopes for resurrection of frozen Neanderthals. Alas, the mechanisms by which reptiles and frogs survive freezing dash those hopes.

Ectotherms living at high latitudes use three alternative strategies for surviving subzero winter temperatures. Some species, such as many Arctic and Antarctic fishes and insects, use glycerol or specific proteins as antifreezes to supercool while avoiding freezing⁷. Others, including most known high-latitude amphibians, snakes and turtles, find locally non-frozen conditions in burrows below the frostline, at the bottom of ponds, or in communal dens. Finally, many terrestrial arthropods and intertidal invertebrates can survive extracellular freezing to as low as -70°C . The last strategy requires that several obvious dangers be overcome: ice crystals can cause physical damage; freezing must be confined to the extracellular fluid; the resulting dehydra-

tion raises solute concentrations and shrinks cells; and circulatory stasis threatens damage through anoxia.

The reptile now shown to survive these dangers is the hatchling Canadian painted turtle *Chrysemys picta marginata* (see figure). It was already known that these turtles hatch in autumn and overwinter in nests in shallow soil that offers little insulation against frost. Storey et al.¹ found by thermocouple recordings that winter nest temperatures repeatedly dropped below zero to as low as -8°C , yet all hatchlings found in nests in the spring were still alive.

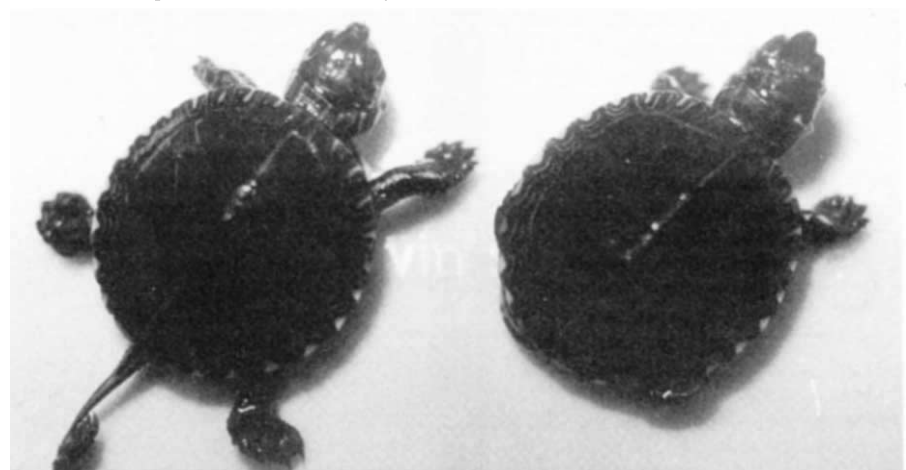
When brought into the laboratory and

cooled, the turtles supercooled to -3.3°C and then froze. Ice accounted for 52–53 per cent of total body water in turtles kept at -4°C . When rewarmed and thawed, those turtles had a strong heart beat, good blood flow, normal breathing, normal head and eye movements, and normal ability to walk and respond to pinches. Thus, the animals must have frozen and thawed repeatedly in their winter nests. But animals frozen in the laboratory to -10.9°C — a temperature below that recorded in their nests — had 59–66 per cent total body water as ice and failed to survive thawing.

Like many freeze-tolerant invertebrates, the turtle and the four freeze-tolerant frogs use low-molecular-mass cryoprotectants to lower their freezing point colligatively, to keep much of their water in liquid form, and to stabilize their proteins. Because glycerol is the most common cryoprotectant in invertebrates and also in one species of frog, it was surprising to find glucose used instead as cryoprotectant at concentrations up to 0.5 molar in the other three frogs, especially as glucose concentrations are so closely regulated in most vertebrates. The turtles provided a further surprise by using a cocktail of glycerol, glucose and amino acids (particularly taurine). The mechanism used by wood frogs is activation of glycogen phosphorylase and consequent production of glucose from liver glycogen.

Another surprising difference between freeze-tolerant vertebrates and invertebrates is that only the latter practise anticipatory cryoprotection: invertebrates gradually synthesize their cryoprotectants before the usual freezing season. In contrast, the initiation of freezing itself stimulates the turtle and the wood frog rapidly to generate raised cryoprotectant concentrations — this process takes as little as a few minutes in the frog species.

Synthesis of cryoprotectants is not the sole adaptation enabling vertebrates and invertebrates to survive freezing. Also involved are adaptations of proteins, membranes and mitochondria; tolerance



Hatchling painted turtles, taken from their nest alive in April after having spent the winter frozen.

J.M. Diamond

of anoxia and very low metabolic rates; and (in invertebrates but not vertebrates) specific ice-nucleating proteins to control ice formation. As extracellular freezing halts blood flow and hence oxygen delivery, for example, the residual metabolism must be supported anoxically. The dependence of freeze-tolerant turtles on anaerobic glycolysis is shown by the decline in glycogen and the accumulation of lactate in most organs during freezing.

In Eurasia during the Ice Ages, many an unlucky Neanderthal and Cromagnon must have succumbed to freezing outdoors, as people still occasionally do today. Imagine the insights to be gained about Palaeolithic culture if such an individual could be brought back to life (see John Marshall's recent News and Views article⁸, for example). Frozen mammoths have been found, and it is possible that the body of a frozen Palaeolithic hunter will be discovered some day as well. Although the discoveries of the first freeze-tolerant vertebrates seem to

bring the dream of thawing out such a specimen tantalizingly closer, on reflection they actually push it even further into the realm of science fiction. Only by a series of specialized adaptations lacking in most animals do the painted turtle and four frog species confine freezing to the extracellular space and so avoid lethal damage. Hence we must continue to reconstruct Palaeolithic culture by traditional, laborious, indirect methods. □

Jared M. Diamond is a professor of Physiology at the University of California Medical School, Los Angeles, California 90024-1751, USA.

1. Storey, K.B. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **85**, 8350–8354 (1988).
2. Storey, K.B. & Storey, J.M. *J. comp. Physiol.* **B155**, 29–36 (1984).
3. Storey, J.M. & Storey, K.B. *J. comp. Physiol.* **B156**, 191–195 (1985).
4. Storey, K.B. *Cryo-letters* **6**, 115–134 (1985).
5. Storey, K.B. & Storey, J.M. *Copeia* (3), 720–726 (1987).
6. Storey, K.B. & Storey, J.M. *Physiol. Rev.* **68**, 27–84 (1988).
7. Raymond, J.A. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **86**, 881–885 (1989).
8. Marshall, J.C. *Nature* **338**, 702–703 (1989).

EVOLUTIONARY BIOLOGY

Asynchronous hatching in birds

William J. Sutherland

MANY species of bird start incubating their eggs before the clutch is complete so that the resulting nestlings differ in age. In one extreme case, a barn owl (*Tyto alba*) was observed to hatch its seven eggs over 15 days. In 1947, David Lack¹ suggested that asynchronous hatching is an adaptation to cope with unpredictable conditions during the raising of the brood. Inequalities in age ensure that, during periods of food shortage, the youngest will starve soon after hatching while the rest will have sufficient food: during periods of abundant food a large brood can be reared. On page 536 of this issue², Magrath describes the first full test and confirmation of Lack's hypothesis.

Magrath studied blackbirds (*Turdus merula*), a species in which hatching asynchrony occurs and the youngest chick grows more slowly and often starves. He manipulated the asynchrony within some broods and provided additional food to some territories. He shows that, during food shortage, asynchronous broods are more productive as brood reduction occurs earlier, and the surviving young are heavier and more likely to survive as juveniles. In broods raised with abundant food, the fledgling survival, fledgling weight and productivity are similar for asynchronous and synchronous broods. Magrath's study shows how asynchronous hatching, coupled with brood reduction, adjusts the brood to the feeding conditions.

In other studies, the process by which brood reduction takes place has been investigated — in some cases it is dramatic

and vicious. In many species of eagles, pelicans, boobies, owls and cranes, the oldest sibling attacks the younger³. In detailed observations of a Verreaux's eagle (*Aquila verreauxi*) nest, Garnett quantified the attacks of a nestling on its younger sibling⁴. The youngest chick continually lost weight and died after three days. In all it was attacked in 38 separate bursts, totalling 1,569 pecks, and was even pulled from under its brooding mother so that the attack could continue.

It has been suggested that this aggression is related to the amount of food available such that in times of food shortage the younger chick is killed. There is clear evidence for this in the blue-footed booby (*Sula nebouxi*) studied by Drummond and colleagues^{5–7}. This species has a modal clutch of two; there is a mean difference of four days between the two chicks. The youngest chicks are fed less frequently and grow more slowly than their siblings. They are more likely to die and this often appears to be a result of aggression by the older sibling. There is no evidence that the junior chicks that die are lighter than the junior chicks that survive. The crucial factor determining the survival of the junior chick is the condition of the senior one. The chicks that died have lighter older siblings which average 20–25 per cent below their potential weight⁵. This implies that brood reduction is caused by attacks of the older offspring; further observations confirmed this⁶ as the amount of aggression towards the younger chick increases during periods of food shortage

and intensifies markedly when the senior chick is about 20–25 per cent below its potential weight.

Other species, such as herons, show similar patterns in brood reduction in relation to the amount of food even though the aggression is not linked to food supply. Mock *et al.*⁸ studied great egrets (*Casmerodius albus*) and great blue herons (*Ardea herodias*) and showed that the aggression is just as high in nests in which the parents bring considerable quantities of food as those in which they bring little. Providing additional food to the egrets similarly does not reduce the fighting rate. However, the survival of the youngest chicks is dependent upon food supply — apparently the ability to withstand the attacks increases if they are better fed⁸.

Thus many studies of asynchronous species show how brood reduction is dependent on food supply, with youngsters being starved through competition or killed by direct attacks: Magrath's work shows how asynchronous hatching and brood reduction can increase productivity. In some species, brood reduction has developed further so that the older chick always kills the younger. Garnett's detailed observations of Verreaux's eagle⁴ seem representative of the species as of more than 200 nests with clutches of two eggs there was only one in which both survived to fledge. The same pattern is shown in other birds of prey and in masked booby (*S. dactylatra*) and brown booby (*S. leucogaster*), in which it is extremely rare for two individuals to be raised^{3,7}.

These species kill their siblings even in the presence of abundant food. In the case of Verreaux's eagle there were five hyrax (*Procavia capensis*) carcasses weighing a total of 5.7 kg in the nest; the mother attempted to feed both birds⁴. There are, however, benefits in producing two eggs, as those nests with two eggs produce a fledgling on 76.4 per cent of occasions compared with only 49.0 per cent of nests with single eggs⁹. In these species it seems that asynchrony and brood reduction is not an insurance against unpredictable feeding conditions but is an insurance against the first egg being infertile, or the embryo or chick failing to survive. □

William J. Sutherland is in the School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK.

1. Lack, D. *Ibis* **89**, 302–352 (1947).
2. Magrath, R.D. *Nature* **339**, 536–538 (1989).
3. Mock, D.W. in *Infanticide: Comparative and Evolutionary Perspectives* (eds Hausfater, G. & Hrdy, S.B.) 3–30 (Aldine, New York, 1984).
4. Garnett, V. *Ostrich* **49**, 57–63 (1978).
5. Drummond, H., Gonzalez, E. & Osorno, J. *Behav. ecol. Sociobiol.* **19**, 365–372 (1986).
6. Drummond, H. & Chavelas, C.G. *Anim. Behav.* **37**, 806–819 (1989).
7. Drummond, H. in *Acta 19 Congressus Internationalis Ornithologica* (ed. Ouellet, H.) (University of Ottawa Press, in the press).
8. Mock, D.W., Lamey, T.C. & Ploger, B.J. *Ecology* **68**, 1760–1772 (1987).
9. Stinson, C.H. *Evolution* **33**, 1219–1225 (1979).

Magnetic fields in superlattices

L. Eaves

THE physics of electron transport in two dimensions is one of the most active fields in solid-state physics today. Perhaps the best-known property of the two-dimensional electron gas is the quantum Hall effect¹. Much of the latest research in this area involves the use of electrostatic gating to modify further the electron motion (see the previous News and Views article² I wrote with F.W. Sheard). In a series of new articles, two West German groups^{3,4} describe the use of this technique to examine the electrical conductivity of a two-dimensional electron gas subjected to the combined effect of a magnetic field and a weak periodic potential. They observe a new oscillatory magnetoresistance effect which has not been observed in the 'natural' periodic potential which arises from the regular atomic

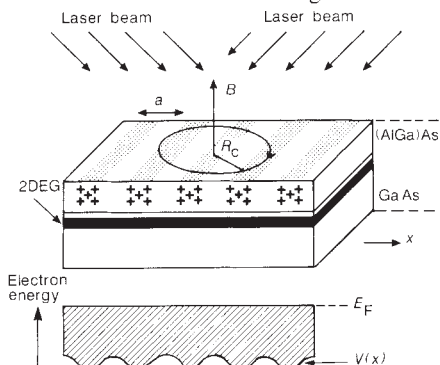


FIG. 1 The arrangement used by von Klitzing's group^{3,4}. Top, two interfering laser beams produce a pattern of bright and dark parallel fringes on the surface of the sample. The bright fringes ionize some of the impurities (+) in the (AlGa)As layer. The resulting periodic space charge variation gives rise to a one-dimensional periodic potential $V(x)$ for the conduction electrons in the 2DEG (bottom). The direction of the applied field B and a cyclotron orbit, radius R_c , are also shown.

arrangement in metallic or semiconductor crystalline lattices.

Advances in semiconductor device fabrication — molecular-beam epitaxy (MBE) and so on — have led to remarkable developments in solid-state physics. In particular, devices built on a sub-microscopic length scale reveal pronounced quantum effects in the electronic motion. The high crystalline perfection and purity achievable with MBE effectively eliminate the scattering processes which can mask these effects. One of the new systems studied, the two-dimensional electron gas (2DEG), can be produced at the interface of a GaAs/(AlGa)As heterojunction. The electronic potential energy perpendicular to the interface has the form of a potential well which confines the electrons to free motion only in the plane of the interface.

Several theoretical studies (see ref. 6) have been made of the quantum mechanics of electrons subjected to both a periodic potential, modulating their linear motion, and a magnetic field applied perpendicular to the 2DEG plane, promoting circular orbital motion. An interesting problem of commensurability arises because there are two fundamental length scales in the system: the period a of the potential and the magnetic length $l_B = (\hbar/eB)^{1/2}$. The ratio $\alpha = a^2/2\pi l_B^2 = \alpha B/(h/e)$ is the number of flux quanta passing through a unit cell. (The quantum of magnetic flux $\phi_0 = h/e$.) The energy level diagram of an electron forms an unusual and striking pattern — referred to as the Hofstadter Butterfly⁷ — when plotted against α . In semiconductor and metal crystals, moving electrons experience a periodic potential owing to the regular arrangement of atoms in the crystal lattice. In this case, the length a corresponds to the lattice constant of the crystal. But because the interatomic separation in crystals is very small, of the order of 2 Å, a magnetic field of around 10⁷ tesla (T) would be required to thread one flux quantum through an atomic unit cell. Such fields are about 300 times larger than those achievable in the laboratory. Hofstadter suggested that the case of $\alpha > 1$ could be investigated by preparing a synthetic two-dimensional periodic potential with a much larger value of a and that the periodic potential could be generated using the electrostatic field effect on which semiconductor transistor action is based.

The two West German groups^{3,4} have used this technique in a heterostructure based on GaAs/(AlGa)As. But instead of using a two-dimensional periodic potential, they investigate electron motion in the presence of a magnetic field and a potential which varies periodically in one direction. The amplitude of the variation is small compared to the kinetic energy of the electrons. In both cases the measurements were carried out at liquid-helium temperatures (below 4.2 K). This reduces the effect of scattering of electrons by atomic vibrations. Winkler *et al.*⁵ produced the one-dimensional superlattice by depositing an array of metallic stripes on the surface of the (AlGa)As layer. The array can be biased with an applied voltage and acts as a gate, as in a field-effect transistor.

Von Klitzing's group^{3,4} in Stuttgart used a different method. They illuminated their cooled sample briefly with a well-defined parallel fringe pattern produced by two interfering laser beams (Fig. 1). The light ionizes some of the impurities in the

(AlGa)As layer, and these remain ionized when the light is switched off. Hence the 2DEG experiences a periodic potential resulting from the stripe-shaped array of impurity space charge. Winkler *et al.* also used a laser interference method in fabricating their metallic gate array, so that the period of their superlattice — $a \approx 0.5 \mu\text{m}$, corresponding to the wavelength of visible light — was roughly the same as that in the Stuttgart experiment.

The magnetoresistance of these devices, measured with magnetic field applied perpendicular to the plane of the 2DEG, includes a new type of oscillatory effect at fields up to about 0.4 T (Fig. 2). The field values at which peaks occur form a series which is periodic in the inverse of the magnetic field and given by

$$2R_c = \frac{4\pi\hbar}{eB\lambda_F} = \left(m + \frac{1}{2}\right)a$$

Here $2R_c$ is the diameter of the cyclotron orbit of an electron at the Fermi energy E_F (see Fig. 1), λ_F is the corresponding de

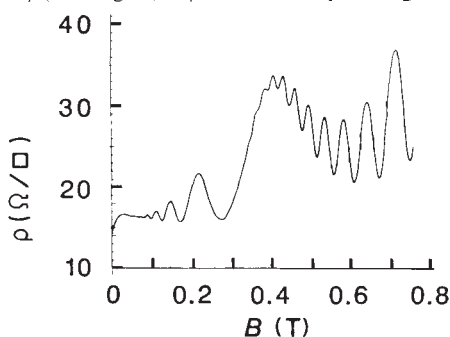


FIG. 2 Oscillatory magnetoresistance pattern for current perpendicular to the interference fringes. The new series occurs for B up to about 0.4 T. The oscillations at higher field are due to the Shubnikov-de Haas effect. The electron density $N_s = 3.16 \times 10^{11} \text{cm}^{-2}$, the electron mobility $\mu = 1.3 \times 10^6 \text{cm}^2 \text{V}^{-1} \text{s}^{-1}$ and the superlattice period $a = 382 \text{nm}$. The measurement was made at $T = 2.2 \text{K}$.

Broglie wavelength, $m = 1, 2, 3, \dots$ and ϕ is a phase factor ($0 < \phi < 1$). At low temperatures, the ballistic path length of the conduction electrons is long enough to allow electrons to make several cyclotron orbits without scattering. The equation relates the three key length scales in the experiment — a , λ_F and the magnetic length l_B . At magnetic fields above about 0.4 T, the new magneto-oscillations are replaced by those arising from the well

1. von Klitzing, K., Dorda, G. & Pepper, M. *Phys. Rev. Lett.* **45**, 494–497 (1980).
2. Sheard, F.W. & Eaves, L.W. *Nature* **333**, 601–602 (1988).
3. Weiss, D., von Klitzing, K., Ploog, K. & Weimann, G. *Proc. Int. Conf. Applications of High Magnetic Fields in Semiconductor Physics 1988* (Springer, in the press).
4. Gerhardt, R.R., Weiss, D. & von Klitzing, K. *Phys. Rev. Lett.* **62**, 1173–1176 (1989).
5. Winkler, R.W., Kotthaus, J.P. & Ploog, K. *Phys. Rev. Lett.* **62**, 1177–1180 (1989).
6. Hofstadter, D.R. *Phys. Rev.* **B14**, 2239–2249 (1976).
7. Hofstadter, D.R. *Gödel, Escher, Bach* (Harvester, 1979).
8. Beenakker, C.W.J. *Phys. Rev. Lett.* **62**, 2020–2023 (1989).

known Shubnikov–de Haas effect (attributable to a different quantum effect). These are also periodic in inverse magnetic field and the period $\Delta(1/B)$ provides an accurate means of measuring the electron density N_s and Fermi wavelength λ_F of the 2DEG: $\Delta(1/B) = 2e/hN_s = e\lambda_F^2/h\pi$.

The two West German groups use a similar theoretical model, based on quantum mechanical perturbation theory, to explain the magneto-oscillations. An analysis by Beenakker⁸ based on classical

mechanics gives a similar result. He shows that the centre of the cyclotron orbit drifts along the direction of the gate stripes when the 'resonance' condition given by the equation is satisfied. This drift motion of the electrons arises from the combined action of forces due to the magnetic field and the electric field of the superlattice. Off resonance, the orbit centre drift is negligible. □

L. Eaves is in the Department of Physics, Nottingham University, Nottingham NG7 2RD, UK.

COSMOLOGY

Explosive assault on Ω

Robert P. Kirshner

THE explosive effect of supernova 1987A on astrophysics vividly reminds us that supernova explosions mark the violent demise of stars, generate 10^{53} erg blasts of neutrinos, forge the elements out of which astronomers and telescopes are made, and produce (perhaps) violently spinning neutron stars. But what are they good for? On page 523 of this issue¹, an international collaboration including workers from Denmark, Australia and the United Kingdom report heroic work to discover and study the most distant supernova yet: an object at an apparent visible-band magnitude of 22.05 — near the limits of detection — in a cluster of galaxies at redshift $z = 0.31$. The authors suggest that using distant supernovae as standard candles with 10^6 times the luminosity of Cepheid variables will allow us to measure the global geometry of space, and to determine the cosmological density parameter Ω .

The technique is straightforward, but tedious. Using the 1.5-m Danish telescope at La Silla, Chile, the Danish part of the team have monitored 60 clusters in the redshift interval $0.2 < z < 0.5$ over a period of two years. Each month, for nine months of the year, they obtain 1-hour CCD (charge-coupled device) exposures and immediately compare the current frame with previous data to look for the subtle brightening that a supernova makes in the image of a distant galaxy. The group has already had one success², detecting an object that might be a type II supernova in a cluster at $z = 0.28$, but the current report is more revealing because a spectrum of this faint new object has been obtained at the Anglo-Australian Telescope.

The spectrum is important, because there are several types of supernovae, and they are sorted by their spectra. The type I supernovae are those with no hydrogen present, in contrast to the type II (such as SN1987A) in which hydrogen is a conspicuous feature. In recent years, it has become clear³⁻⁵ that the type I objects must be further divided into two classes —

type Ia, which show a distinct absorption attributed to Si II at a wavelength of 6,150 Å, and type Ib which lack that feature. This difference is not a mere spectroscopic detail, but is correlated with different energy output at optical, infrared and radio frequencies and probably reflects different stellar progenitors for the two types. The distant supernova measured last August has a noisy spectrum because it is so faint, and it is superposed on the light of the galaxy that contained its progenitor, but the spectrum looks remarkably like those of type Ia supernovae nearer us, once the effects of the redshift are taken into account.

The cosmological use of the type Ia objects hinges on their homogeneity⁶. The theoretical reason is that the events are likely to arise from the subsonic thermonuclear burning of carbon–oxygen white dwarfs exactly at the Chandrasekhar mass — those only just stable against gravitational collapse. Such a distinct and well-defined process may be expected to lead to a well-defined light output. The empirical reason is that a carefully selected set of good observations for type Ia supernovae shows a small (less than 0.2 mag) scatter in the magnitude at maximum^{7,8}.

With such a small intrinsic scatter, the measurement of distant supernovae of this type can lead to a measure of the cosmological density Ω by the classical methods of the redshift–magnitude diagram. In euclidean space, the brightness of a distant standard candle drops off as the square of the distance, but gravitational space curvature induced by the presence of matter can produce measurable deviations from this relation. It is this prospect that Norgaard-Nielsen *et al.* hold out in this issue¹. The apparent magnitude for a type Ia object at $z = 0.3$ differs by 0.16 mag between $\Omega = 0.0$ and the $\Omega =$ exactly 1.0 so beloved of inflation theorists (Ω is normalized to the critical density needed to halt, eventually, the Universe's expansion). If the scatter in the observations is of the same order, then in a few years, observa-

tions of a handful of supernovae in distant clusters might begin to narrow the allowed range of Ω . The advantage of this approach is that the type Ia events being observed in distant clusters may be the same events as we can observe nearby: if that is exactly true, then no corrections for evolution need to be applied.

Although it is appealing to think that supernovae might lead us out of an age of ignorance and belief into an era of measurement and understanding, two observational issues need to be carefully studied before too much faith is placed in this promising approach. First, the homogeneity of the type Ia events is a matter of observation, not of faith, and there are recent spectroscopic clues that show small, but real, differences among members of this class⁹. If there is a corresponding photometric variation, then the task of determining the difference between $\Omega = 0$ and $\Omega = 1$ from the Hubble diagram may move from the formidable to the insuperable.

Second, we need to build confidence that the supernovae observed at high redshift are really the same as the supernovae observed nearby. Although the spectrum and the light curve obtained for the supernova at $z = 0.31$ resemble those of nearby supernovae, some subtle selection effect may make the supernovae found in distant clusters a little different from the local sample. A cautious way to approach this problem is to populate the Hubble diagram at intermediate redshift from $z = 0.03$ to $z = 0.3$, and to provide empirical evidence that type Ia supernovae remain superlative standard candles from our own Galaxy out to the edge of the observable Universe. The authors of the article in this issue¹ suggest that because the sought-for effect would be larger at $z = 0.5$, the best course is to press on to the highest possible redshift. However, another approach might be to shore up our knowledge of supernovae from the local neighbourhood through intermediate redshift to be certain that any observed effect comes from space curvature and not from a changing population of supernovae. □

Robert P. Kirshner is at the Harvard-Smithsonian Center for Astrophysics, Cambridge, Massachusetts 02138, USA.

1. Norgaard-Nielsen, H.U. *et al.* *Nature* **339**, 523–525 (1989).
2. Hansen, L., Norgaard-Nielsen, H.U., Jorgensen, H.E., Ellis, R.S. & Couch, W.J. *Astr. Astrophys.* **221**, L9–L11 (1989).
3. Porter, A.C. & Filippenko, A.V. *Astr. J.* **93**, 1372 (1987).
4. Uomoto, A.K. & Kirshner, R.P. *Astr. Astrophys.* **149**, L7 (1985).
5. Schlegel, E.M. & Kirshner, R.P. *Astr. J.* (in the press).
6. Branch, D.M. in *The Extragalactic Distance Scale* (eds van den Bergh, S. & Pritchet, C.J.) 146–156 (Astronomical Society of the Pacific, 1988).
7. Leibundgut, B. thesis, Univ. Basel (1989).
8. Cadonau, R., Sandage, A., Tammann, G.A. in *Supernovae as Distance Indicators* (ed. Bartel, N.) 151–165 (Springer, Berlin, 1985).
9. Branch, D.M., Drucker, W. & Jeffery, D.J. *Astrophys. J.* **330**, L117–L118 (1988).

Do tolerant T cells exist?

Daniel L. Mueller

THE immune system has evolved to seek out and destroy invading pathogens. Discriminating between 'self' and 'non-self' is vital; cells which recognize self must behave differently from those which do not if the damaging consequences of autoimmunity are to be avoided. For T cells this might be achieved in three different ways. First, T cells bearing receptors that recognize self antigens could be physically eliminated as they develop in the thymus (clonal deletion). Second, self-reactive T cells that do mature and enter the periphery could be functionally inhibited by other T cells (suppression/immunoregulation). Third, when self-reactive T cells encounter the antigen they recognize, they could be functionally inactivated, instead of being activated, to effect an immune response (clonal anergy).

How far each of these mechanisms contributes to developing and maintaining self-tolerance remains unclear. There is solid evidence for a physical loss of potentially self-reactive clones during the course of development within the thymus. By contrast, direct evidence for antigen-specific suppression or regulation of mature T cells has been difficult to obtain. The existence of functionally inactivated T cells *in vivo* has also suffered from lack of direct proof. On page 541 of this issue¹, however, Rammensee *et al.* show that tolerant T cells really do exist. They demonstrate that clonal anergy, rather than clonal deletion, can account for antigen-specific tolerance induced in adult mice. Added to other evidence, their results suggest a role for clonal anergy in the maintenance of natural self-tolerance *in vivo*. Similarities with previously described models of induced non-responsiveness *in vitro* point to common biochemical principles that may underlie all three modes of tolerance.

The antigen on which the experiments of Rammensee *et al.* are focused is the so-called *Mls* determinant. The products of the *Mls* genes are polymorphic proteins of mice known only from the vigorous mixed-lymphocyte reactions (MLR) they evoke between cells from mice identical at the major histocompatibility complex (MHC) but bearing different *Mls* alleles (see ref. 2 for a review). The vigour of these responses *in vitro* is now known to result from the very high frequency of *Mls*-reactive T cells: for example, most T cells whose receptors contain a β -chain encoded by either *V β 8.1* or *V β 6* gene segments recognize *Mls*²; and in mice bearing the *Mls*^a allele, tolerance is ensured by the deletion of T cells expressing these *V β* segments during thymic ontogeny.

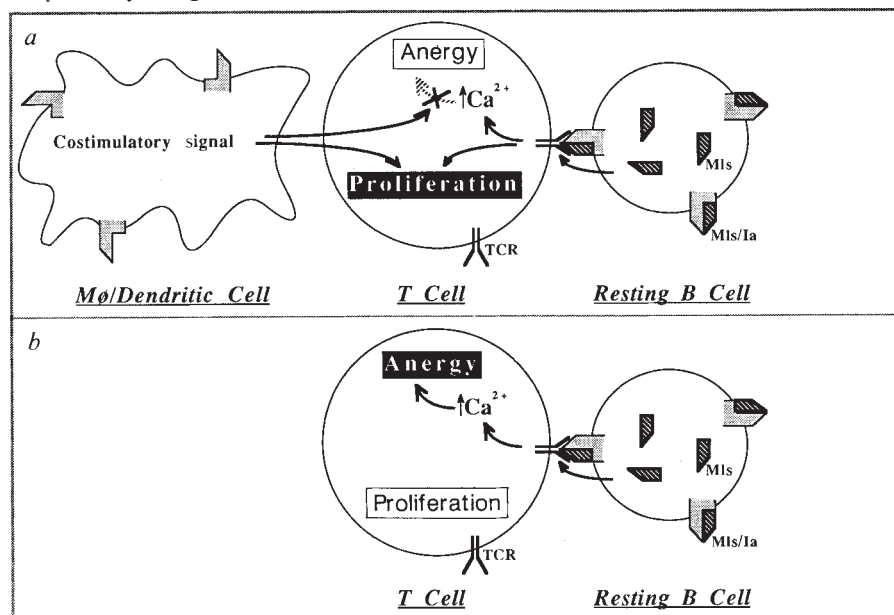
Given the exceptionally strong response

in vitro to cells mismatched at *Mls*, one would expect a correspondingly strong response *in vivo*. Yet *Mls*^b mice do not reject MHC-identical *Mls*^a skin grafts. In fact, the mixed-lymphocyte reaction of *Mls*^b mice to *Mls*^a cells can be abrogated by prior injection of adult *Mls*^a cells into the *Mls*^b mice³. Thus it seems that exposure of *Mls*^b mice to *Mls*^a determinants *in vivo* leads not to priming of an immune response but to the induction of tolerance.

Rammensee *et al.*¹ now show how this unresponsiveness is achieved. They find no evidence for a physical deletion of the *Mls*-reactive clones in these animals; *V β 6*⁺ CD4⁺ T cells remain highly represented within the lymph nodes. This is consistent with the results of Qin *et al.*⁴, who recently found, while examining the induction of transplantation tolerance to minor histocompatibility antigens, that tolerance to

produce interleukin-2. These experiments are the first to demonstrate that clonal anergy can account for *in vivo* tolerance to an antigen.

The contribution of functional T-cell inactivation to the development of natural self-tolerance cannot be assessed from these data alone. Recent experiments by Marrack *et al.*⁵ demonstrate that radiation bone-marrow chimaeras and transgenic mice that express a particular MHC antigen only on the stromal (non-haematopoietic) elements of the thymus fail to delete physically T lymphocytes that are potentially reactive to that self-antigen. Despite this, these animals demonstrate tolerance to the antigen in two ways. First, T-cell reactivity to the antigen in MLR is significantly reduced; and second, these animals demonstrate no autoimmunity against their thymic tissue. Marrack *et al.* are not convinced that clonal anergy exists here because T-cell hybridomas produced from fusion with these self-specific lymphocytes are normally functional. On the other hand, it is not clear that



Proliferation versus anergy. *a*, *In vitro*; *b*, *in vivo*. TCR, T-cell receptor.

Mls^a could develop in mice despite the continued presence of normal numbers of *V β 6*⁺ T cells in their lymph nodes. Rammensee *et al.* show that responses against MHC-disparate spleen cells (presumably by the *V β 6*⁺ T-cell population) in primary MLR are unaffected by the state of unresponsiveness to *Mls*^a, and find in cell-mixing experiments that no *Mls*-specific suppression exists. Remarkably, the authors demonstrate the mature *Mls*-responsive T cells to be functionally inactivated. The density of T-cell receptors is normal and interleukin-2-receptor expression is induced on these cells following stimulation; thus, receptor occupancy and some form of signalling can take place. Despite this, the T cells fail to proliferate because of an inability to

functionally inactivated T cells would retain their phenotype on fusion with a tumour line. This is now testable with fusions of the *V β 6*⁺ population described by Rammensee *et al.*

The lack of clonal deletion of T cells specific for antigens not expressed on thymic haematopoietic elements raises the possibility of a role for clonal anergy in various experimental tolerance models. Chicken chimaeras that bear embryonic thymic stromal elements transplanted from quails do not reject either this tissue or a quail limb⁶. Allogeneic *Xenopus* (amphibian) chimaeras created by fusion of the anterior portion of one strain with the posterior portion of another fail to come apart⁷. Transgenic mice expressing a particular MHC-antigen only on pancreatic

β -cells tolerate this tissue and fail to respond to the MHC-antigen in MLR⁸. In none of these models would clonal deletion be expected to play the major role in developing tolerance, as the particular antigens involved are not expressed on haematopoietic elements within the thymus. Clonal anergy seems a possible mechanism, although immunoregulation or a suppression of autoreactivity could also be responsible.

The findings of Rammensee *et al.* have striking similarities to an *in vitro* murine model of T-cell unresponsiveness first described by Jenkins and Schwartz⁹. In this system, interleukin-2-driven proliferation by type I CD4⁺ T-cell clones responding to peptide antigen and syngeneic splenic antigen-presenting cells (APC) is blocked if the APC population is pretreated with a chemical fixative. Rather than stimulating proliferation, these cells, in the presence of peptide antigen, induce the T cells into a state of proliferative unresponsiveness. This state of unresponsiveness appears to be entirely analogous to that seen in the $V_{\beta}6^{+}$ population following treatment of the mice with Mls^{a+} spleen cells — T-cell receptor expression is intact, yet interleukin-2 production and proliferation cannot be induced in response to normal APC and antigen.

Comparisons to this *in vitro* model of unresponsiveness may provide clues to the mechanism of the induction of T-cell Mls -tolerance *in vivo*. Two points in particular are worth considering. First, the induction of unresponsiveness *in vitro* is dependent on the mobilization of intracellular calcium in response to T-cell receptor occupancy¹⁰. Second, an accessory cell-derived costimulatory signal, delivered at the time of the calcium flux, is capable of blocking the induction of unresponsiveness, and is also responsible for synergizing with signals generated via the T-cell receptor to induce the production of interleukin-2 and T-cell proliferation¹¹. A failure to deliver this costimulatory signal during the $V_{\beta}6^{+}$ T-cell response to the transferred Mls^{a+} spleen cells *in vivo* could be important for the induction of the tolerant state.

The Mls^{a} antigen appears to be expressed primarily on B cells². Resting B cells represent a poor source of costimulatory signals, as judged by their poor ability to augment the proliferation of T cells responding to antigen presented by fixed APC¹¹, and by demonstrations of their poor stimulatory capacity in primary MLR despite reasonable expression of MHC molecules¹². These cells are, however, capable of inducing a T-cell proliferative response to the Mls^{a} antigen *in vitro*. Costimulatory signals must certainly be provided by nearby splenic macrophages or dendritic cells to account for the strong proliferation in these cultures. The delivery of this second signal would also

block the induction of unresponsiveness in the culture (*a* in the figure). The response seems to be different on recognition of the Mls^{a+} B cell *in vivo*. The development of the anergic state suggests that little costimulatory signal is delivered at the time of occupancy of T-cell receptors, and the effects of the calcium flux are dominant (*b* in the figure). Why these T cells fail to receive the second signal from neighbouring accessory cells (as would seem to be the case *in vitro*) is unclear. The ability of Batista and Bloom¹³ as well as Miller *et al.*¹⁴ to induce tolerance to several antigens following the intravenous delivery of antigen-coupled or haptenated spleen cells suggests that the route of administration and presumably the tissue compartment within which the recognition of Mls^{a} occurs are critical variables.

The biochemical basis of proliferative unresponsiveness *in vitro*, namely increases in intracellular calcium concentration in the absence of a costimulatory signal, could also represent a common biochemical principle that underlies each of the models of T-cell tolerance. Mobilization of calcium in thymocytes undergoing negative selection seems to be the trigger that induces apoptosis and the deletion of these self-reactive clones¹⁵. The recent demonstration that CD8⁺-suppressor T cells can induce clonal anergy in CD4⁺ T cells responding to *Mycobacterium leprae* antigens *in vitro* suggests that suppressor T cells may operate by interfering with the delivery of a costimulatory signal to other antigen-stimulated T cells¹⁶. Such common themes may turn out to be important to our understanding of T-cell tolerance as the questions of deletion versus suppression versus anergy are answered. □

Daniel L. Mueller is in the Laboratory of Cellular and Molecular Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA.

- Rammensee, H.-G., Kroschewski, R. & Frangoulis, B. *Nature* **339**, 541–544 (1989).
- Janeway, C.A. Jr *et al.* *Immun. Rev.* **107**, 61–88 (1989).
- Jacobson, J., Lilliehook, B. & Blomgren, H. *Scand. J. Immun.* **4**, 181–191 (1975).
- Qin, S., Cobbold, S., Benjamin, R. & Waldmann, H. *J. exp. Med.* **169**, 779–794 (1989).
- Marrack, P. *et al.* *Cell* **53**, 627–634 (1988).
- Ohke, H., Martin, C., Corbel, C., Coltey, M. & Le Douarin, N.M. *Science* **237**, 1032–1035 (1987).
- Flajnik, M.F., Du Pasquier, L. & Cohen, N. *Eur. J. Immun.* **15**, 540–547 (1985).
- Lo, D. *et al.* *Cell* **53**, 159–168 (1988).
- Jenkins, M.K. & Schwartz, R.H. *J. exp. Med.* **165**, 302–319 (1987).
- Jenkins, M.K., Pardoll, D. M., Mizuguchi, J., Chused, T. M. & Schwartz, R.H. *Proc. natn. Acad. Sci. U.S.A.* **84**, 5409–5414 (1987).
- Jenkins, M.K., Ashwell, J.D. & Schwartz, R.H. *J. Immun.* **140**, 3324–3330 (1988).
- Minami, M., Shreffler, D.C. & Cowing, C. *J. Immun.* **124**, 1314–1321 (1980).
- Batista, J.R. & Bloom, B.R. *Nature* **212**, 156–157 (1966).
- Miller, S.D., Sy, M.-S. & Clamen, H.N. *Eur. J. Immun.* **7**, 165–170 (1977).
- Smith, C.A., Williams, G.T., Kingston, R., Jenkinson, E.J. & Owen, J.J.T. *Nature* **337**, 181–184 (1989).
- Salgame, P., Modlin, R. & Bloom, B.R. *Int. Immun.* (in the press).

Coca-coda

COCAINE addiction is a growing problem in many cities of the Western world. Daedalus now plans to counter it with a novel strategy based on a curious episode in nineteenth-century imperialism. In those innocent days, cocaine was a valued medical drug; and to capture the market for it, the Dutch tried to transplant coca plants from South America to their colonies in Indonesia. The plants grew quite well: but, bafflingly, yielded almost no cocaine.

What was going on? Daedalus reckons that the plants were infected by some wild plant virus from the local forests. Not all such viruses are deadly; some merely cause small metabolic changes. So Daedalus advocates an expedition to Indonesia to identify this virus, the plants it naturally infects, and the insects which spread it. These could then be covertly transported to South America, and released around the illicit coca plantations.

The results would be most intriguing. Modern *Erythroxylon coca* is a highly cultivated plant, bred over centuries to produce far more cocaine than is good for it. Freed from this pointless biochemical burden, the infected coca plantations would flourish as never before, but would enigmatically fail to yield any cocaine. Apoplectic drug barons and mafia bosses would suddenly find themselves making a loss on their investment. The surprise would be even greater if the virus turned out not to suppress alkaloid production, but switched it instead to some alternative alkaloid like strychnine. Once well entrenched in the local vegetation, the virus would be impossible to eradicate. All concerned would have to find another racket, or even (if all else failed) an honest trade.

This subtle principle can clearly be extended. Once identified, the plant virus might be mutated or genetically engineered to switch the alkaloid metabolism of other pharmacologically active plants. Suitable variants could subvert the opium poppy fields of Asia or the many illicit cannabis plantations around the world. Anti-smoking fanatics would love to release a nicotine-suppressing virus onto the world's tobacco fields. But Daedalus himself wants to challenge the hypocrisy of society's wine snobs with a virus that suppresses the alcohol synthesis of wine yeast, but leaves the complexities of fermentation otherwise untouched. The resulting wine will of course retain all the much vaunted subtleties of bouquet and palate that *bon viveurs* insist are the reason for their interest and devotion; but it will no longer make them drunk. The price which this new vintage commands, and the degree of dedication with which it is bought, laid down, argued and enthused about, compared with other vintages, and finally sampled, will be most revealing.

David Jones

Cold fusion doubts and controls

SIR—Much premature excitement has been generated by the claims and counter-claims for induced cold fusion. There is apparently no consensus (*Nature* **338**, 616; 1989) on the central questions regarding the amount of excess heat produced, the evidence for nuclear reactions and the vexed question of whether the detection of reaction fragments could be consistent with the claims for excess heat generation.

For many years we have been investigating primordial nucleosynthesis in a hot Big-Bang cosmology. The measured abundances of the light isotopes of hydrogen, helium and lithium place stringent restrictions on a number of physical parameters and are sensitive to deuterium-deuterium reaction rates. We have investigated by how much these reaction rates can be increased by a single low-energy resonance, consistent with deuterium-deuterium laboratory data, without destroying the ability to provide a consistent description of primordial light-isotope abundances.

Armed with these maximum reaction rates, we can estimate the density of deuterium packing in the electrodes of one of the cold-fusion experiments in order to generate excess heat at the rate of watts, as reported in some experiments. Our result is 10^7 mol cm⁻³. We conclude that more heat has been generated in the media coverage of this subject than by deuterium fusion in electrolytic cells.

J. M. IRVINE
S. RILEY

Department of Theoretical Physics,
Schuster Laboratory,
The University,
Manchester M13 9PL, UK

SIR—Fleischmann, Pons and Hawkins' claimed¹ detection of neutrons and tritium from a postulated d-d reaction is not yet supported by their account of the controls used to eliminate other sources of both neutrons and tritium. The control on neutron background reported in ref. 1—a series of measurement with the deuterium-containing cell removed—is no control on neutrons arising from photodisintegration of deuterons or other materials in the cell by ambient γ -rays. Such a background could also be a problem with the reported neutron flux of Jones *et al.*². A check on that background would be to determine whether the neutrons appeared and disappeared when the power to the electrochemical cell was switched on and off.

The tritium measurement is even more problematic; one must have controls against the possibility of neutron capture creating tritium by the reactions $D(n,\gamma)T$ and ${}^6Li(n,\alpha)T$. Because tritium decay cannot be turned off, supplementary

experiments may be needed to show that tritium does not grow into a vessel freshly filled with tritium-free D and 6Li by either of the postulated reactions. Tritium is also known to exchange quickly with other hydrogen isotopes, is widely used in commerce³, and is a durable contaminant of laboratories which have been exposed to tritium in various forms.

Note that repetition of the experiments in refs 1 and 2 with deuterium replaced by hydrogen is not a control on the suggested sources of neutrons or tritium. A negative result with tritium-free hydrogen is therefore moot with respect to claims of cold fusion based on neutron or tritium detection in a D- 6Li mixture.

LAWRENCE CRANBERG

TDN Inc., Austin,
Texas 78764, USA

1. Fleischmann, M., Pons, B. & Hawkins, M. J. *electroanal. Chem.* **261**, 301–308 (1989).
2. Jones, S. E. *et al.* *Nature* **338**, 737–740 (1989).
3. Moghissi, A. A. *et al.* (eds) *Radioactivity in Consumer Products* (US Nuclear Regulatory Commission, Washington, DC, 1978).

Radiolysis of water

SIR—Baverstock and Cundall¹ suggest that the energy imparted directly and locally by ionizing radiation to a DNA molecule may manifest itself as broken bonds at a distance because, between the time of energy deposition and bond breaking, the energy is converted to a soliton, which may comprise a wave of coupled vibrational and electronic excitation travelling along the molecule.

In an electron-nuclear double-resonance study² of the crystalline malonic acid-urea 1:1 adduct (again a hydrogen-bonded structure) irradiated at room temperature, a signal was observed which was ascribed to a solitonic wave. This wave is excited and sustained by the presence of double-well hydrogen bonds in the crystal. The crystal structure consists of planar layers held together by a complex network of hydrogen bonds within the bond planes.

If we extend these ideas from the one- and two-dimensional structures above to a three-dimensional case, the classic example of such a hydrogen-bonded structure is liquid water. The initial behaviour of water irradiated at room temperature exhibits certain anomalies³, which defy explanation if we base our ideas about energy deposition from high-energy radiation on gas-phase interactions, because this approach takes insufficient account of liquid structure.

In particular, in the *ab initio* modelling of product yields, the stochastic method of kinetic calculations⁴, when used in conjunction with a well-known method of deriving the distribution⁵ in spur energy, requires (in an acid system) the unit radius

for H or e_{aq}⁻ to be as low as 0.44 nm and to be 0.19 nm for OH, in order to obtain the observed results for $G(H_2)$ and $G(H_2O_2)$ (where $G(M)$ is the number of molecules of M formed per 100 eV of radiation energy absorbed). In contrast, the deterministic method requires more credible values of 2.3 and 0.75 nm, respectively³. This is partly because, for the more abundant smaller spur sizes, the stochastic method gives a greater ratio of combination for unlike relative to like pairs of radicals.

This difficulty would be mainly obviated if the typical entity decomposed by radiation were a group of five to six molecules⁶, part of the 'flickering cluster' structure of liquid water⁷, in which the secondary electrons in the lower part of the energy distribution (below, say, 50 eV) give rise to a solitonic wave whose energy could be dissipated in these clusters in non-dissociative processes, but in which higher electron energies would cause dissociation and ionization. Such a model would remove the proportionality previously assumed between the energy per spur and the number of ionizations per spur (taken, for example, to be 6 ionizations per 100 eV in ref. 5, or 12 radicals per 100 eV in ref. 4), irrespective of the amount of energy imparted. To obtain the correct average ionization yield (about one ion pair per 20 eV), the energy transferred in packets larger than a threshold level of about 50 eV would need to be more efficient in producing ionizations, but the use of this modified distribution in the number of ionizations per spur would then give results close to those for a set of uniform spurs of 5–6 ionizations per spur, for which the stochastic and deterministic models give similar results⁸. In addition, the resulting, more realistic, unit spur radii of 2.3 nm for aquated electrons and 0.75 nm for hydroxyl radicals would give rise to agreement between prediction and observation in the effects of scavengers and of linear energy transfer, and also in time-resolved studies^{9,10} as there is reasonable agreement in all these areas when the deterministic method is used^{3,6}.

An alternative stochastic approach¹¹ is showing interesting possibilities in some of these areas, especially in predicting scavenger effects at early times and time-resolved yields of e_{aq}⁻ for fast electrons⁹ and for 3-MeV protons¹⁰—but the role of normalization at 10 ps in the *ab initio* calculation of ref. 11 was not adequately discussed.

In cases where the first-order structure of the water is broken down, such as during high-temperature (300 °C) γ -radiolysis (A.J. Elliot *et al.*, personal communication) or for densely ionizing radiation at room temperature¹², more product formation is observed than can be accounted for by normal diffusion-controlled kinetics. This may be because the lower-energy secondary electrons

interact with a 'flickering cluster' size distribution which has a greater abundance of smaller clusters, thereby causing decomposition.

W.G. BURNS

Chemistry Division,
Harwell Laboratory,
Didcot, Oxfordshire OX11 0RA, UK

1. Baverstock, K.F. & Cundall, R.B. *Nature* **332**, 312–313 (1988).
2. Colligiani, A., Pinzino, C. & Angelone, R. *Molec. Phys.* **62**, 97–107 (1987).
3. Burns, W.G., Sims, H.E. & Goodall, J.A.B. *Int. J. Radiat. Phys. Chem.* **23**, 143–163 (1983).
4. Clifford, P., Green, N.J.B., Pilling, M.J., Pimblott, S.H. & Burns, W.G. *Int. J. Radiat. Phys. Chem.* **30**, 125–132 (1987).
5. Mozumder, A. & Magee, J.L. *J. chem. Phys.* **45**, 3332–3341 (1966).
6. Schwarz, H.A. *J. phys. Chem.* **75**, 1928 (1969).
7. Erdelyi-Gruz, T. *Transport Phenomena in Aqueous Solutions* (Akadémiai Kiadó, Budapest, 1974) transl. Ruff, I., rev. Eged, Y. (Adam Hilger, London).
8. Green, N.J.B., Pilling, M.J., Clifford, P. & Burns, W.G. *J. chem. Soc. Faraday Trans.* **80**(1), 1313–1322 (1984).
9. Jonah, C.D. & Miller, J.R. *J. phys. Chem.* **81**, 1974 (1977).
10. Burns, W.G., May, R., Buxton, G.V. & Wilkinson-Tough, G.S. *J. chem. Soc. Faraday Trans.* **77**(1), 1543 (1981).
11. Brenner, D.J. *Int. J. Radiat. Phys. Chem.* **32**, 157–162 (1988).
12. Burns, W.G. & Sims, H.E. *J. chem. Soc. Faraday Trans.* **77**(1), 2803–2813.

Unmasking N-CAM

SIR—Using heterokaryon analysis, Pavlath *et al.*¹ recently demonstrated the restricted expression of the antigen reactive with the monoclonal antibody 5.1H11 (ref. 2) to localized regions of the plasma membrane defined by the position of active nuclei. We wish to point out that the antigen recognized by 5.1H11 is human N-CAM, an adhesion molecule originally discovered in neural cells.

Since its initial identification, the molecular nature of the 5.1H11 antigen has remained undefined due to the detergent-sensitive nature of the reactive epitope. We have, however, recently cloned human complementary DNAs encoding the major alternatively spliced N-CAM isoforms of skeletal muscle^{3–5}. In the course of expression studies of these human N-CAMs we examined 3T3 fibroblasts stably transfected with human N-CAM cDNAs for reactivity with 5.1H11 using indirect immunofluorescence. Whereas parental mouse cells were

unreactive with 5.1H11, all transfected cell lines were strongly reactive. The figure shows an example. As the antibody reacts with a species-specific epitope⁶, artefactual or indirect activation of a previously silent mouse gene is excluded. Thus, the 5.1H11 antigen is unequivocally identified as human N-CAM. Similar results were found with many independently derived clones and reactivity was also found with different isoforms of N-CAM, including transmembrane, glycosylphosphatidylinositol-linked and secreted forms. Therefore, previous data indicating that a diffusible *trans*-activator regulates 5.1H11 expression in muscle cells⁶ in fact relates to regulation of the gene for N-CAM.

N-CAM expression increases in association with myotube formation and there is also a switch from a splicing pathway generating a transmembrane N-CAM to one producing a glycosylphosphatidylinositol-linked isoform^{3,4}. Isoform-specific antibodies or cDNAs might thus be used to determine which isoforms are produced in the heterokaryon model and which is localized over active nuclei.

FRANK S. WALSH
GEORGE DICKSON
STEPHEN E. MOORE
C. HOWARD BARTON

Department of Neurochemistry,
Institute of Neurology,
Queen Square, London WC1N 3BG, UK

1. Pavlath, G.K., Rich, K., Webster, S.G. & Blau, H.M. *Nature* **337**, 570–573 (1989).
2. Walsh, F.S. & Ritter, M.A. *Nature* **289**, 60–64 (1981).
3. Dickson, J.G. *et al. Cell* **50**, 1119–1130 (1987).
4. Barton, C.H. *et al. Development* **104**, 165–173 (1988).
5. Gower, H.J. *et al. Cell* **55**, 955–964 (1988).
6. Blau, H.M. *et al. Science* **230**, 758–766 (1985).

Sequence identity

SIR—We wish to draw attention to the complete identity between the C-terminal sequence of the cytosolic calcium-binding protein (CaBP) p14 and the N terminus of two small neutrophil-immobilizing-factor proteins, NIF-1 and NIF-2.

The two associated cytosolic CaBPs¹, p8 and p14 (also known as CF antigen², LI molecule³, or MRP8 and MRP14 (ref. 4)) are members of a larger family of CaBPs that may be involved in intracellular signal transduction. All the proteins have a molecular mass of about 10,000 with the exception of p14 which has a longer C-terminal sequence following the second calcium-binding domain, prompting speculation that this region of the molecule performs a unique function⁵. By searching the OWL database⁶ with the sequence of p14 from position 89–108, we found that it is identical to the N-terminal 20 amino-acid residues of NIF-1 and NIF-2 (refs. 7 and 8), with the exception of an extra alanine residue at the N terminus of NIF-2. We suggest that residues 89–108 of

p14 contain a sequence that performs an extra-cellular function of immobilizing neutrophils engaged in chemotaxis. It is also possible, however, that the common sequence of p14 and NIF serves some other function and that the neutrophil-immobilizing activity actually resides in the C-terminal half of the NIF peptides.

A further possibility is that the NIF peptides are formed by proteolytic cleavage of p14 after the tryptophan residue at position 88. If this is so, then two features of the NIF peptides require further explanation. First, cleavage of p14 at position 88 would yield a C-terminal peptide of 26 amino acids in length, whereas NIF-1 and NIF-2 are predicted from their amino-acid compositions to contain 41 and 38 amino acids respectively⁸. But recalculation of the amino-acid composition of NIF-1 for a peptide of 26 residues produces a composition almost identical to that of the C-terminal tail of p14 with the exception of 0.4 mol mol⁻¹ of cysteic acid. A similar analysis of NIF-2 could not be accounted for.

It was originally reported that p8 and p14 are associated with a separate macrophage-migration inhibitory factor (MIF)⁹. We now suggest that a neutrophil immobilizing activity is part of the p14 protein itself. Our previous work indicates that the p8, p14 complex (or some component part) is released by myeloid cells during their interaction with vascular endothelium¹. Therefore p14 could be involved in immobilizing myeloid cells at the endothelial surface both during an inflammatory response and during the normal exudation of myeloid cells into tissues.

PAUL FREEMONT

NANCY HOGG

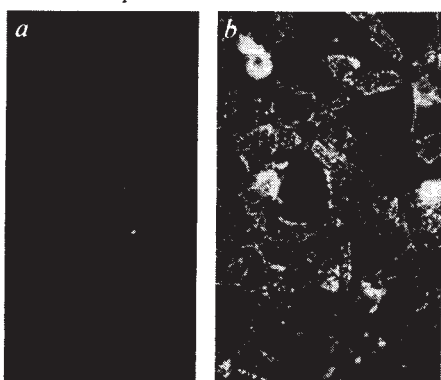
JONATHAN EDGEWORTH

Imperial Cancer Research Fund,
Lincoln's Inn Fields,
London WC2A 3PX, UK.

1. Hogg, N., Allen, C. & Edgeworth, J. *Eur. J. Immun.* (in the press).
2. Dorin, J.R. *et al. Nature* **326**, 614–617 (1987).
3. Dale, I., Fagerol, M.K. & Naesgaard, I. *Eur. J. Biochem.* **134**, 1–6 (1983).
4. Odink, K. *et al. Nature* **330**, 80–82 (1987).
5. Lagasse, E. & Clerc, R.G. *Molec. cell. Biol.* **8**, 2402–2410 (1988).
6. Akrigg, D. *et al. Nature* **335**, 745–746 (1988).
7. Goetzl, E.J. & Austen, K.F. *J. exp. Med.* **136**, 1564–1580 (1972).
8. Watt, K.W.K., Brightman, I.L. & Goetzl, E.J. *Immunology* **48**, 79–86 (1983).
9. Burmeister, G., Tarcsay, L. & Sorg, C. *Immunobiology* **171**, 461–474 (1986).

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. They need not arise out of anything published in *Nature*. In any case, priority will be given to letters of less than 500 words and five references. □



Indirect immunofluorescence staining using monoclonal antibody 5.1H11 with a, 3T3 cells or b, 3T3 cells transfected with a human N-CAM cDNA.

Disentangling the bank

John Lawton

Perspectives in Ecological Theory. Edited by Jonathan Roughgarden, Robert M. May and Simon A. Levin. Princeton University Press: 1989. Pp. 394. Hbk \$60; pbk \$22.50.

Too many field ecologists hate theory. How can the tangled bank, in all its glorious, muddy complexity, be reduced to a set of antiseptic differential equations? To be fair, it isn't certain that it can; but to presume that we cannot hope to find general rules for the way nature works is to reduce every ecological interaction to a special case, to be marvelled at and then added to the vast rummage box of wonderful anecdotes and special cases. Certainly, ecological systems are complex, sometimes bewilderingly so; a humble English meadow may be home to more like a thousand species than a hundred. Nevertheless, there are sufficient regularities and patterns in nature to suggest that there are rules and regulations underlying the tangle.

The present book has its origins in a growing sense of alarm and frustration about the apparently widening rift between theoreticians and field workers in ecology, and between pattern seekers and the gee whiz school of wonderful anecdotes. In an attempt to forge closer links between both sides of the divide, some 40 ecologists representing the broad sweep of the subject met in Asilomar, California, in 1987. The fruits of their labours are displayed in these *Perspectives*.

There are eight main sections, spanning the core of theoretical ecology, including population dynamics, species interactions, ecology and evolution, community and ecosystem processes, and resource management. With one exception, each section contains two review chapters followed by a discussion chapter in which one of the participants attempts to draw together the threads of what politicians often refer to as "frank and wide-ranging discussions". In the main this format works very well.

For those who require convincing that empirical ecologists need theory to guide their work, Kareiva's chapter, entitled "Renewing the Dialog between Theory and Experiments in Population Ecology", is essential reading; the few short paragraphs on pages 78–80 ought to be engraved in the hearts and minds of all aspiring ecology graduates, and those of their teachers. If this fails to move them, then Powell's simple observation that "we simply cannot afford to wait for the data to pour in to determine what the future may bring for our planet" should do the trick. Developing theoretical ecology is not just an absorbing intellectual exercise. It isn't even a matter of life and death; it's much

more important than that.

The book has many strengths. Not least, it provides a gold mine of interesting and important questions, sufficient to keep armies of graduate students profitably employed for several years. Perhaps more importantly it defines an agenda for research initiatives to be undertaken by the entire community. For example, a theme repeated in at least eight chapters is the problem of how ecologists should deal



Tracking back — a female turtle returns to sea after laying her eggs. The picture is reproduced from the chapter on population dynamics in the new, fifth edition of *Biology*, by Helena Curtis and N. Sue Barnes, published by Worth.

with and model nested hierarchies of spatial and temporal scales. The question emerges in such disparate areas as plant physiological ecology (Chapter 1, by Gross), behavioural ecology and resource acquisition (Chapter 2, by Pulliam), ecosystem structure and function (Chapter 16, by Levin) and forest simulators (Chapter 17, by Horn, Shugart and Urban). Two contributions are devoted entirely to the problem, one general and wide-ranging by O'Neill and the other more specific, on physical and biological scales of variability in aquatic ecosystems, by Powell. Although no clear answers emerge about how we should deal with the problem, it is a sign of the growing maturity of ecology that such wildly different subject areas should converge on essentially the same set of abstract, general questions.

From the many other key issues raised by the book, I will select just two. In the discussion on ecosystem structure and function (Chapter 18), Schlesinger asks whether "some species, or some ecosystems, are more important to global function than others". And in Chapter 20, Pimm and Gilpin review the role of

theoretical ecology in conservation biology. As species and entire ecosystems vanish before the flood-tide of humanity, I can think of no more pressing fields of research.

Of course, there are weaknesses. Not all the chapters are as thoughtful, insightful and new as they could and should have been. Indeed, one or two are little more than skimpy rehashes of rather tired themes; fortunately, they are few.

I also think two things are lacking. First, there is little attempt, anywhere in the book, to ask what are the main patterns and regularities in nature that ecologists wish to explain. Tilman (in the discussion to the section on population dynamics and species interactions, Chapter 6) and Cody (in the discussion on structure and assembly of communities, Chapter 15) both emphasize the urgent need to document major patterns, a view that I entirely endorse. But it is not a message that most of the writers have acted upon, except perhaps implicitly. Cohen (Chapter 13) is a notable exception, providing a masterly account of patterns in food webs, together with a unifying hypothesis to explain them.

Second, the book contains a sampling of theoretical ecology, not a comprehensive review. Perhaps it was never intended to do more than select some of the high ground in the centre of the range of possibilities. Nevertheless, it is a pity that some intriguing theoretical questions have not found a place. Many are new, or poorly worked — for example questions about why there are more small species than large species in the world; whether there are hard limits to global diversity; and what determines the size, shape and limits of species' geographical ranges. Many of the issues that form the core of the book have been around for half a century and more. It is good to see real progress being made towards answering them, at last; but it would have been even better to see clearer signposts to the next mountain range.

If *Perspectives in Ecological Theory* succeeds, as I believe it will, in interesting more field workers to take an interest in theory, and more theoreticians to attempt to explain what is really going on out there, rather than what they imagine is going on, and to couch their models in testable form, it will be a valuable contribution to the literature. Whether we like it or not, ecologists have to start giving hard answers to some very hard questions. Without a sound theoretical base for our subject, we will be unable to meet the challenges of the next decade and beyond. Revelling in the complexity of the tangled bank is simply not an option any more. □

John Lawton is Director of the Centre for Population Biology and Professor of Community Ecology at Imperial College, Silwood Park, Ascot, Berkshire, SL5 7PY, UK.

Prognosis for diagnosis

Philip Mortimer

HIV Detection by Genetic Engineering Methods. Edited by Paul A. Luciw and Kathelyn Sue Steimer. *Marcel Dekker*: 1989. Pp. 301. \$99.75 (North America); \$119.50 (elsewhere).

ANYONE who doubts the scope and power of modern biotechnology should look through this book. It shows how, from a standing start in 1984, academic and commercial laboratories in the United States have raised HIV diagnostics from crude art to high science.

A cursory inspection of the contents reveals how busy the research and development laboratories have been and how remarkable their achievements. A more attentive read, even by a beginner, would lead to a broad understanding of the techniques of antibody testing, antigen detection, *in situ* hybridization and gene amplification as they are being

applied to HIV diagnosis. The various authors have placed genetic engineering in the broad context of laboratory diagnosis, and the book illustrates very well how the new approaches are interacting with the old. What is being so rapidly achieved for HIV will very soon be applied in other diagnostic areas and, as the final chapter on HTLV I suggests, this process has already begun. Diagnostic tests will become safer, cheaper, more accurate, more discriminating and more versatile.

The popular conception of genetic engineering is that it will revolutionize therapeutics and immunization: the reality is that diagnostics will be transformed first. Existing tests will be refined and (as has recently been shown for non-A non-B hepatitis) new diagnoses offered. The work on HIV exemplifies this process and here is a volume that conveniently records what is happening. The book will have a short shelf life, but it is an important and thoroughly worthwhile read. □

Philip Mortimer is Director of the PHLS Virus Reference Laboratory, Central Public Health Laboratory, Colindale Avenue, London NW9 5HT, UK.

On coherence

G. H. C. New

Lasers. By Peter W. Milonni and Joseph H. Eberly. *Wiley*: 1988. Pp. 731. £35.85, \$49.95.

IN RAPIDLY developing areas of research, it is only to be expected that textbooks will be few and far between. The people who, in less frenetic times, would be writing them, instead burn the midnight oil in their laboratories. Laser physics is a good example. In barely 30 years, it has grown from nothing into a vast field of endeavour that now impinges on innumerable aspects of science and technology. And although there are a few excellent books on the subject, lecturers giving courses on lasers know only too well how difficult it often is to find a suitable text for their students.

Milonni and Eberly's new book will solve the problem for some people; as a supporting text for an introductory course at the first-year graduate level, it is ideal. The strength of the book lies in its breadth rather than its depth, for the authors' intention is to describe the broad foundations of the subject rather than taking readers into uncharted territory.

Following an elementary introduction to lasers, there are eight chapters concerned with basic radiation theory. All the topics one would expect are covered, including classical dispersion and absorption theory, the Einstein A and B coefficients, the Schrödinger, Maxwell and

Bloch equations, rate equations and so forth. Chapter 4, dealing with the energy-level structure of atoms, molecules and solids (including semiconductors), is particularly nice, while Chapter 9, on photons and photon statistics, might not be found in less comprehensive works.

Basic laser theory is treated in Chapters 10 and 11, and Chapter 12 goes on to deal with laser dynamics and Q-switching, with a brief mention of mode-locking; there are even listings of BASIC and FORTRAN programs so that enthusiastic readers can solve the Wagner-Lengyel equations on their personal computers. Chapter 14 applies the principles set out earlier to dye lasers, gas and gas dynamic lasers, chemical lasers, free-electron lasers and semiconductor lasers. And still to come are chapters on resonator theory, coherence and laser applications, finishing with two chapters on nonlinear optics. Between them, these two chapters take up 90 of the 730 pages in the book; one cannot get too far in nonlinear optics in that space. This is the dilemma that the authors have faced and which has determined the character of the book they have written.

This, then, is a highly successful attempt to cover the foundations of laser physics and nonlinear optics in a single volume. On narrower fronts, other books on laser physics may perhaps be more appropriate. But for the task that the authors have set themselves, it is difficult to see how they could have done a better job. □

G. H. C. New is a Professor in the Department of Physics, Imperial College, Prince Consort Road, London SW7 2AZ, UK.

From Noah on

Hugh Torrens

Catastrophic Episodes in Earth History. By Claude C. Albritton, Jr. *Chapman & Hall*: 1989. Pp. 221. Hbk £27.50, \$29.95; pbk £12.95.

THE subject tackled in Claude Albritton's book is a minefield, one made all the more dangerous by the recent 'catastrophe' of judging scholarship by volume. Albritton starts with an historical survey of disaster and catastrophe from Noah to now — from the Flood, to Werner's German neptunism or Hutton's Scottish plutonism, and beyond. We learn that Werner's lectures drew "hundreds of students from all over Europe", but are not told that far more were attracted to Edinburgh from all over the world. Moreover, students in Edinburgh were not indoctrinated, but could make up their own minds.

The new catastrophism of meteorite impacts, and the 'resulting' mass extinctions and other catastrophe scenarios, are discussed next. Here is a lucid, up-to-date survey of these heated issues, aimed at undergraduates and the many others who are equally bewildered. Because it is well-nigh impossible to write an historical account of such a continuing debate, however, these sections are less satisfactory as history. (I wish historians would notice the existence of Rene Gallant's *Bombarded Earth: An Essay on the Geological and Biological Effects of Huge Meteorite Impacts*; that book was published in 1964, years before the advent of the Alvarez hypothesis of mass extinctions.)

Albritton finally puts the new catastrophism into proper perspective and explains recent ideas which led to it and to dinosaurianism. He takes a sensible, on the fence, stance as to possible causes of mass extinction. One problem in the debate has been the lack of quantification of 'catastrophes'. A Yorkshire vicar, arguing with Lyell in the 1840s, postulated gradual rates of local sedimentary deposition — but at 900 feet (275 m) a month! Only in a British climate could one see such rates as non-catastrophic; thus my review copy arrived during a snow storm after one of the mildest European winters on record. The same day I learnt of Claude Albritton's death and we must mourn a true interdisciplinarian. His last book provides a fitting requiem. □

Hugh Torrens is a Senior Lecturer in the Department of Geology, University of Keele, Staffordshire ST5 5BG, UK.

● Thoughts on a different sort of catastrophe are the subject of Spencer R. Weart's *Nuclear Fear: A History of Images*. The book, now in paperback, was reviewed in *Nature* 335, 775 (1988). Publisher is Harvard University Press, price is \$14.95, £11.95.

RNA-catalysed synthesis of complementary-strand RNA

Jennifer A. Doudna & Jack W. Szostak

Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA

The *Tetrahymena* ribozyme can splice together multiple oligonucleotides aligned on a template strand to yield a fully complementary product strand. This reaction demonstrates the feasibility of RNA-catalysed RNA replications.

THE recent identification of several catalytically active RNA molecules¹⁻⁶ has led to extensive speculation concerning the role of RNA in the origin of life⁷⁻¹¹. A self-replicating RNA or related polynucleotide is thought to be the key intermediate in the evolution of living systems from prebiotic chemicals. The identification or design of an RNA replicase would therefore be helpful in establishing the feasibility of this pathway, and might also allow the construction of simple proto-cells in the laboratory. The protozoan *Tetrahymena* self-splicing intron^{1,2,6} is of particular interest in this respect because of the variety of phosphoester transfer reactions that it can catalyse. The activities of this ribozyme, demonstrated in a series of experiments by Cech and coworkers, include that of a ribonuclease, phosphotransferase, acid phosphatase and RNA restriction endonuclease¹²⁻¹⁴.

An enzyme capable of catalysing transesterification reactions on RNA substrates is potentially capable of catalysing RNA polymerization. Indeed, Zaug and Cech¹² and Been and Cech¹⁵ have shown that the intron will catalyse limited polymerization of ribonucleotides onto a short primer annealed to a sequence within the intron. But the primer could be extended only to a maximum of ~15 nucleotides, and the nucleotides were added beyond the end of the template¹⁵. These experiments, although demonstrating the ability of the *Tetrahymena* ribozyme to polymerize RNA, also pointed out the problems to be overcome before replicase activity could be approached: first, the template must be on a separate molecule from the replicase; second, the polymerase activity must be template directed; and third, the nucleotide specificity that ensures accurate splicing must be overcome to allow copying of an arbitrary template sequence. Here we show that these three problems can be overcome, and that a modified version of the *Tetrahymena* ribozyme can catalyse the formation of an RNA molecule complementary to a template strand.

Catalysis on an independent template

The initial step in self-splicing (Fig. 1) is the attack of the 3' hydroxyl of a free guanosine on a specific phosphate in a stem-loop of the intron referred to as P1. A phosphoester transfer reaction occurs in which the exon-intron junction is cleaved and the G becomes attached to the 5' end of the intron. P1 is the first of a series of base-paired stems in the intron; the 3' strand of P1 (the internal guide sequence) has been used as a primer-binding site for primer elongation experiments^{12,15}. To overcome the requirement that a replicase act on a template that is a separate molecule, we synthesized the isolated P1 stem-loop (Fig. 1a) by transcription of a synthetic oligonucleotide, and designed a modified version of the ribozyme extending from stem P2 to P9 (Fig. 1b). We found that this enzyme RNA catalyses the site-specific attack of guanosine on the isolated P1 stem (Fig. 2), but that the K_m for free P1 was very high (>0.1 mM). This weak interaction probably reflects the fact that

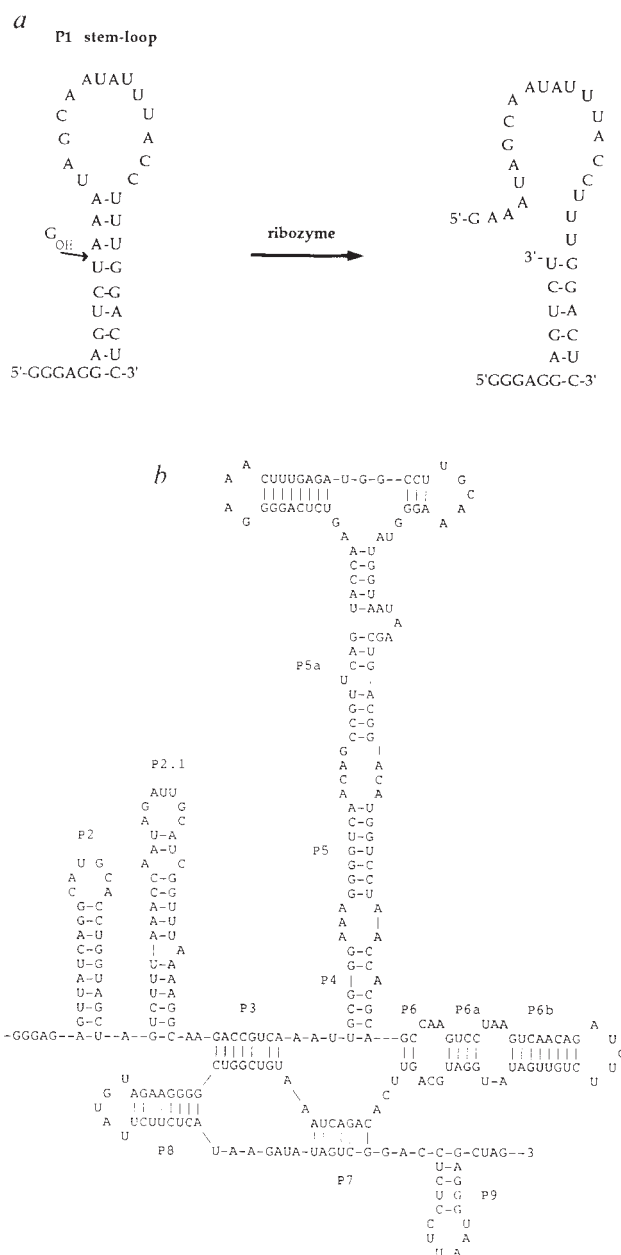


FIG. 1 a, Cleavage of an independent P1 substrate. The diagram illustrates the structure of the substrate used in this experiment and the guanosine cleavage reaction catalysed by the ribozyme. P1 is normally the first stem-loop of the intron, containing the 5' exon-intron junction and the internal guide sequence; in our experiments it is a separate molecule. b, Sequence and secondary structure of the modified *Tetrahymena* intron used in this paper. This molecule is missing the P1 stem-loop, the 3' stem-loops P9.1 and P9.2, and the 3' intron-exon junction. RNA was synthesized by T7 RNA polymerase runoff transcription of plasmid pJD1100 digested with the restriction enzyme *NheI*, and purified either by electrophoresis on a 6% polyacrylamide, 7M urea gel, or by chromatography on a Sephadex G-75 spin column. In either case, the RNA was extracted with phenol, precipitated with ethanol, and stored in distilled water at -80°C .

there are few sequence or size requirements for recognition of P1 by the core intron¹⁶⁻¹⁸.

As a first step towards the assembly of oligonucleotides on an external template, we synthesized two RNA oligonucleotides that correspond to the products of guanosine attack on P1 (Fig. 3a). The sequences are different from that of wild-type P1, apart from the U·G base pair at the site of the guanosine cleavage reaction. The two oligonucleotides anneal to form a complex that consists of an RNA primer annealed to a partial hairpin RNA, with a guanosine residue extruded from the helix at the gap. Incubation of this complex with the modified ribozyme resulted in the extremely efficient regeneration of intact P1, with the release of free guanosine (Fig. 3b). This is the equivalent of the reverse of the reaction shown in Fig. 1a. The oligonucleotide ligation reaction was essentially complete after one hour, with about 250 turnovers of substrate per enzyme.

Spermidine overcomes sequence specificity

We have previously shown, in a different system, that only the wobble base pairs U·G and C·A in the intact P1 stem allow efficient guanosine attack¹⁸. To evaluate the base-pair requirements for the reverse reaction (oligonucleotide ligation), we synthesized four primers ending in either A, C, G or U, and four partial hairpins with either A, C, G or U opposite the last base of the primer (Fig. 4a). The 16 primer-hairpin combinations were then tested for ligation by the ribozyme (Fig. 4b). In contrast to the cleavage reaction, only the U·G and C·G base combinations allow efficient ligation. Ligation occurs to a lesser extent with the A·G combination.

The same set of primer-template combinations was tested for ligation by the intron under various reaction conditions. Because we have hypothesized that base-pair geometry at the reaction site is critical¹⁸, it seemed possible that subtle changes in the structure of either P1 or the ribozyme might allow ligation to occur with other base-pair combinations. One of the conditions tested was the addition of 5 mM spermidine to the reaction (Fig. 4c). This led to the efficient ligation of all the substrate complexes with either Watson-Crick or wobble (U·G, C·A or A·G) base pairs at the ligation junction. A set of 10 related polyamines was tested for similar effects, and none was as effective as

spermidine, although putrescine and spermine work to a lesser extent. Spermidine overcomes the inherent sequence specificity of the ribozyme, and should, in principle, allow the production of a faithful (that is, entirely Watson-Crick) copy of a template.

Template-directed oligonucleotide ligation

Previous work had shown that the loop of the P1 stem did not seem to be important in the guanosine-attack reaction. This suggested that we could simply dispense with the P1 loop in the reverse reaction, and try to ligate together two short oligonucleotides aligned on a longer third oligonucleotide which would act as a template. Of the two shorter oligonucleotides, the 5' one is referred to as the primer, and the 3' one as the ligator. The ligator oligonucleotide begins with a guanosine residue that is not paired with the template, just as in the P1 regeneration experiments described above. Several different primer, ligator and template combinations were designed: three substrates contained a U·G at the ligation junction, whereas four others contained the Watson-Crick base pairs at the junction. In each case, the ribozyme catalysed the ligation of the primer and the ligator in a template-dependent manner (Fig. 5). The number of turnovers per enzyme molecule ranged from 100 in a 15-min reaction for the U·G substrates, to 30-100 in 60 min for the C·G, G·C and A·U substrates, to a low of 5 in 60 min for a U·A substrate. The ligated products were characterized by T1 digestion; in all cases the expected fragments were observed. These experiments show that the ribozyme can catalyse oligonucleotide ligation which is independent of sequence.

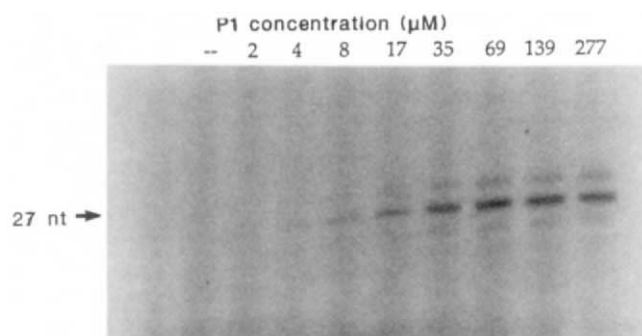


FIG. 2 Cleavage of a P1-like substrate by the modified *Tetrahymena* ribozyme. P1 RNA was prepared by T7 transcription of a synthetic oligodeoxynucleotide¹⁹, followed by purification on a 20% acrylamide, 7 M urea gel. Enzyme reactions contained 10 mM NH_4Cl , 20 mM MgCl_2 , 30 mM Tris-HCl pH 7.4, 1 mM aurin trichloroacetic acid, 0.2 μM enzyme (prepared as in Fig. 1b), 200 μM [α -³²P]GTP, and the indicated concentration of P1 RNA, in a 5- μl reaction volume. Reactions were incubated at 58 °C for 20 min, then stopped by the addition of an equal volume of 90% formamide, 10 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.2% bromophenol blue and xylene cyanol. Aliquots were electrophoresed on a 15% acrylamide, 7 M urea gel; the gel was then dried and autoradiographed. Lanes (left to right): no enzyme, no substrate, increasing substrate concentration. The major band is the expected product of the reaction (see text and Fig. 1a). The minor bands are due to 3' heterogeneity of the substrate, as determined by 5'-end-labelling with polynucleotide kinase.

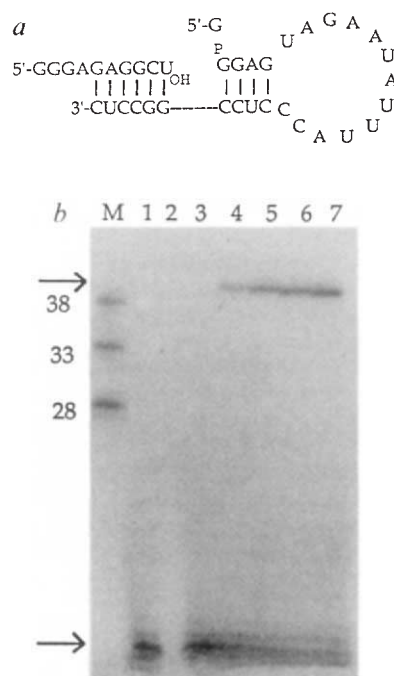


FIG. 3 a, Substrate ribo-oligonucleotides for P1 regeneration. RNAs corresponding to the products of a P1 cleavage reaction were prepared and purified as described above, so that the enzymatic reversal of the cleavage reaction could be followed. The 5' primer oligonucleotide was internally labelled by including 0.5 μCi per μl of [α -³²P]GTP in the transcription reaction. b, Time course of P1 regeneration. Reactions contained 10 mM NH_4Cl , 20 mM MgCl_2 , 30 mM Tris HCl pH 7.4, 1 mM aurin trichloroacetic acid, 0.2 μM enzyme and 50 μM of each RNA substrate oligonucleotide in a total volume of 5 μl . Reactions were incubated at 58 °C, and stopped by the addition of formamide/dye solution as above. Aliquots were electrophoresed on a 20% acrylamide, 7 M urea gel; the gel was dried and autoradiographed. Lane M, DNA size markers; lane 1, no template oligonucleotide, 60 min reaction; lane 2, no primer oligonucleotide, 60 min reaction; lanes 3-7, complete reactions incubated for 0, 15, 30, 45 and 60 min, respectively. Bottom arrow, labelled primer; top arrow, ligated product.

Ligation of aligned multiple oligonucleotides

To see if the ligation reaction explored above could be used to assemble a longer complementary-strand RNA, three sets of templates and complementary oligonucleotides were made (Fig. 6a). In the first set, four identical oligonucleotides align on one long template, with U·G base pairs at each ligation junction. With equimolar template and ribozyme in the reaction, products corresponding to the ligation of two, three or all four oligonucleotides are visible within 5 min of incubation (Fig. 6b). The ratio of full-length to shorter products increases over time, with

concomitant release of free GMP. With smaller amounts of enzyme, the reaction proceeds more slowly.

We considered the possibility that the longer products in the above reaction were generated by a ligation of two oligonucleotides, dissociation from the template and subsequent reannealing to the template with the original ligator in the first position on the template. Longer oligonucleotides could then form by repeated cycles of a reaction very similar to two-oligonucleotide ligation. This model was eliminated by using a template on which four different oligonucleotides could anneal (Fig. 6a), so that each ligation reaction would have to take place at a different position on the template. The course of the reaction was similar to the first multiple ligation (Fig. 6c). The sequence of the full-length product was confirmed by dideoxy sequencing with reverse transcriptase.

In the two experiments described above, the product strand is not fully complementary to the template strand, because of the U·G base pairs at the ligation junctions. We therefore synthesized a template on which five small RNAs could anneal, such that each of the four ligation junctions consisted of a

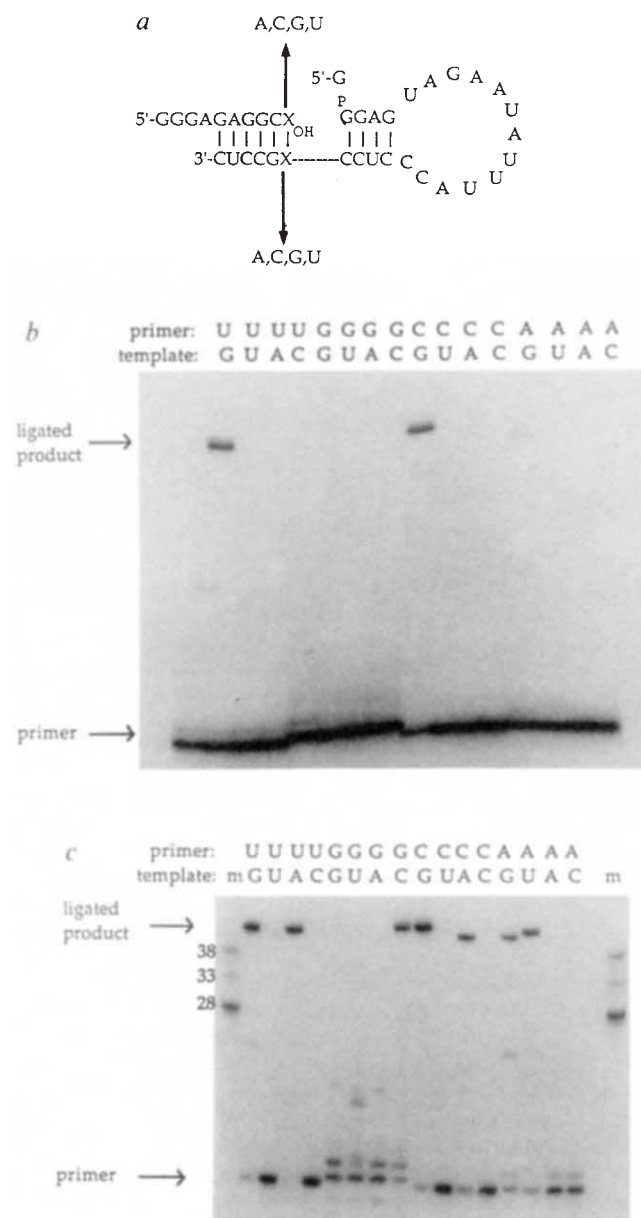


FIG. 4 Effect of the base pair at the ligation junction on P1 regeneration. *a*, We synthesized four primers and four templates, differing at the indicated position. These sets of oligonucleotides can be combined to yield P1 regeneration substrates with any base-base combination at the ligation junction. *b*, P1 regeneration reactions. Reactions were as described above, except that 20 μ M oligonucleotides were used, and incubations were at 58 °C for 60 min. Each reaction contained the indicated base-base combination at the ligation junction. The reaction works well only with the U·G and C·G combinations. *c*, Effect of spermidine on P1 regeneration. Reactions were as above, except for the addition of 5 mM spermidine, which allows U·G, C·A, A·G and all Watson-Crick combinations to ligate efficiently. Ligated products with different sequences have slightly different mobilities.

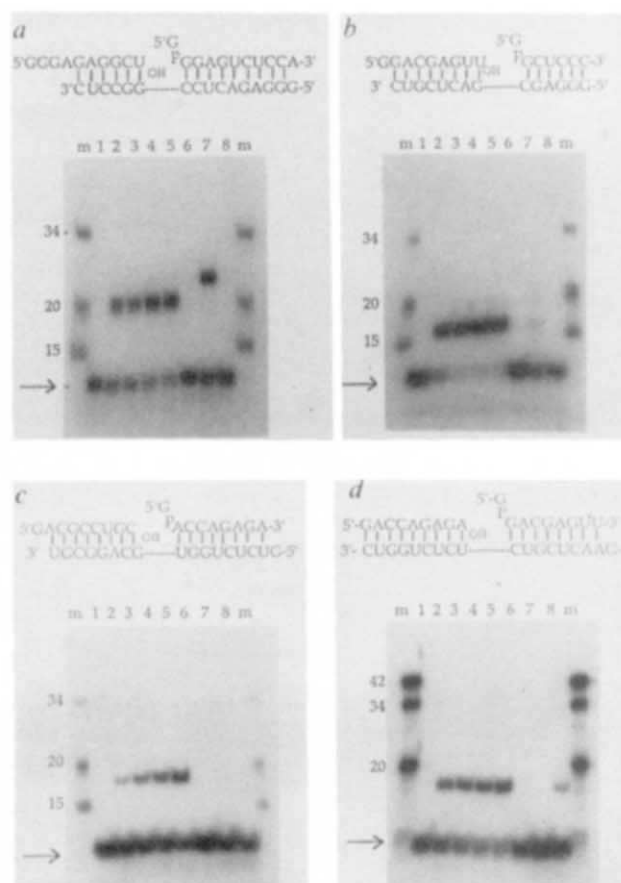


FIG. 5 Template-directed oligonucleotide ligations. Four examples of the ligation of two oligonucleotides aligned by a template oligonucleotide are shown. RNAs were prepared and purified as described above, except that the primer RNAs were labelled with [γ - 32 P]ATP and polynucleotide kinase. Reaction conditions were as in Fig. 4, with 0.2 μ M enzyme, 20 μ M oligonucleotides and 5 mM spermidine. Reactions were stopped by the addition of 3 vol 9.3 M urea, 33 mM EDTA. To denature product template complexes completely, reaction mixes were separated by electrophoresis on a 20% polyacrylamide, 90% formamide gel. Lanes M, RNA oligonucleotide size markers; lanes 1–5, time course, 0, 15, 30, 45 and 60 min, respectively; lanes 6, no template oligonucleotide, 60-min reaction; lanes 7, no ligase oligonucleotide, 60 min reaction; lanes 8, no enzyme (except *a*); lane 7, no enzyme; lane 8, no ligase (except *a*), 60 min reaction. In each case, the arrow represents the position of the primer. *a*, *b*, Substrate complexes with a U·G base pair at the ligation junction; *c*, *d*, complexes with Watson-Crick base pairs at the junction.

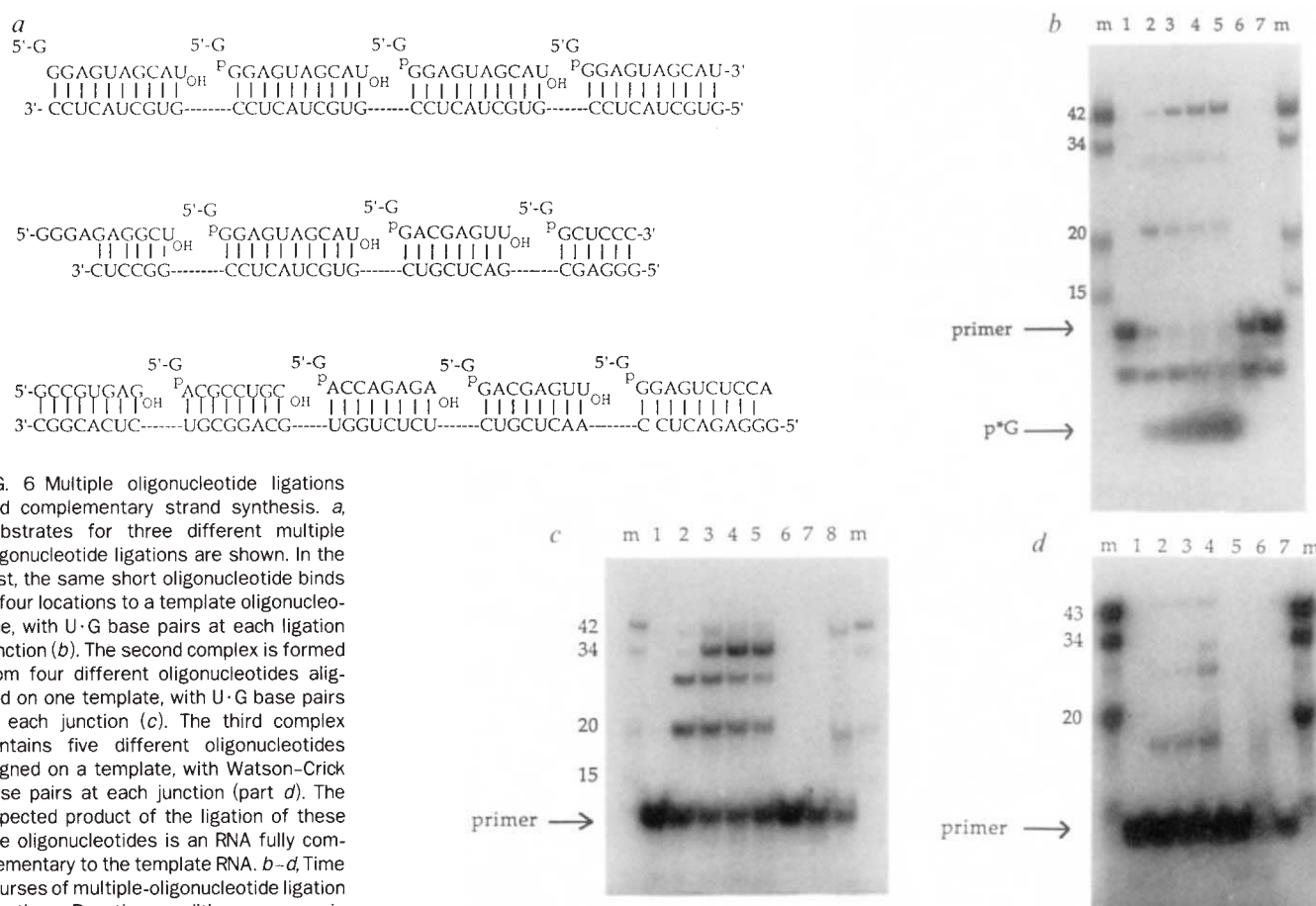


FIG. 6 Multiple oligonucleotide ligations and complementary strand synthesis. *a*, Substrates for three different multiple oligonucleotide ligations are shown. In the first, the same short oligonucleotide binds in four locations to a template oligonucleotide, with U·G base pairs at each ligation junction (*b*). The second complex is formed from four different oligonucleotides aligned on one template, with U·G base pairs at each junction (*c*). The third complex contains five different oligonucleotides aligned on a template, with Watson-Crick base pairs at each junction (part *d*). The expected product of the ligation of these five oligonucleotides is an RNA fully complementary to the template RNA. *b*–*d*, Time courses of multiple-oligonucleotide ligation reactions. Reaction conditions were as in Fig. 5, except that both enzyme and substrate were at 5 μ M. Aliquots of each reaction were electrophoresed on 20% polyacrylamide gels prepared in 90% formamide to promote complete denaturation of the samples. Lanes marked 'm' contain end-labelled RNA size markers. *b*, Ligation of four oligonucleotides shown in Fig. 6*a*, top. Lanes 1–5, complete reactions incubated for 0, 5, 30, 60 and 90 min, respectively; lane 6, no template, 90 min; lane 7, no enzyme, 90 min. *c*, Ligation of four oligonucleotides shown in Fig. 6*a*, middle. Lanes 1–7, as in part *b*; lane 8, no second oligonucleotide. *d*, Ligation of five oligonucleotides shown in Fig. 6*a*, bottom. Lanes 1–4, 0, 30, 60 and 120 min, respectively; lane 5, no enzyme, 120 min; lane 6, no template, 120 min; lane 7, no second oligonucleotide, 120 min.

different Watson-Crick base pair. We observed ~5% ligation to full-length product, with larger amounts of shorter products also accumulating (Fig. 6*d*). The reaction was completely dependent on template and enzyme.

Discussion

The ligation reaction that we have characterized differs from all previously described reactions catalysed by group I introns in that base pairing of substrate RNAs to the intron itself is not required. For example, in the polymerase experiments reported by Cech *et al.*¹⁵, the primer to be extended is bound to the internal guide sequence, which is itself covalently joined to the catalytic domain of the intron. In the experiments described here, the template that aligns the oligonucleotides to be ligated is a separate RNA molecule that does not interact with the enzyme by base pairing. One consequence of this difference is much weaker interaction of our substrates with the enzyme. We think that this weakened interaction is the simplest explanation for the rapid turnover we see in the ligation reaction; after

ligation, the double-stranded product and the single guanosine residue are rapidly released from the enzyme, allowing another cycle of catalysis to begin.

In principle, a ligation reaction similar to that described here could have been used by a crude primordial replicase, and could be used now in the design of an RNA replicase. The rapid turnover that we see is important in this respect, because many catalytic cycles will be required for the replication of a long template. Our experiments show that a wide range of template sequences can be used, and the addition of spermidine to the reaction allows the formation of a fully complementary product. But several refinements will be necessary to achieve autocatalytic replication in the laboratory. The efficiency of full-length strand synthesis must be increased, the possible deleterious effects of template secondary structure must be explored and overcome, and a way to separate product and template strands must be found. These problems can now be approached experimentally; their solution should allow the design of a self-replicating RNA molecule. □

Received 14 April; accepted 8 May 1989.

1. Cech, T. R., Zaug, A. J. & Grobowski, P. J. *Cell* **27**, 487–496 (1981).
2. Kruger, K. *et al.* *Cell* **31**, 147–157 (1982).
3. Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N. & Altman, S. *Cell* **35**, 849–857 (1983).
4. Van der Veen, R. *et al.* *Cell* **44**, 225–234 (1986).
5. Garriga, G. & Lambowitz, A. M. *Cell* **39**, 631–641 (1984).
6. Cech, T. R. *Science* **236**, 1532–1539 (1987).
7. Sharp, P. A. *Cell* **42**, 397–400 (1985).
8. Darnell, J. E. & Doolittle, W. F. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1271–1275 (1986).
9. Gilbert, W. *Nature* **319**, 618 (1986).

10. Cech, T. R. *Proc. natn. Acad. Sci. U.S.A.* **83**, 4360–4363 (1986).
11. Joyce, G. F. *Nature* **338**, 217–224 (1989).
12. Zaug, A. J. & Cech, T. R. *Science* **231**, 470–475 (1986).
13. Zaug, A. J., Been, M. D. & Cech, T. R. *Nature* **324**, 429–433 (1986).
14. Zaug, A. J. & Cech, T. R. *Biochemistry* **25**, 4478–4482 (1986).
15. Been, M. D. & Cech, T. R. *Science* **239**, 1412–1416 (1988).
16. Waring, R. B., Towner, P., Minter, S. J. & Davies, R. W. *Nature* **321**, 133–139 (1986).
17. Been, M. D. *et al.* *Cold Spring Harb. Symp. quant. Biol.* **LII**, 147–157 (1987).
18. Doudna, J. A., Cormack, B. P. & Szostak, J. W. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
19. Milligan, J. F., Groebe, D. R., Witherell, G. W. & Uhlenbeck, O. C. *Nucleic Acids Res.* **15**, 8783 (1987).

The discovery of a type Ia supernova at a redshift of 0.31

Hans U. Nørgaard-Nielsen*, Leif Hansen†, Henning E. Jørgensen†, Alfonso Aragón Salamanca‡, Richard S. Ellis‡ & Warrick J. Couch§

* Danish Space Research Institute, Lundtoftevej 7, DK-2800 Lyngby, Denmark

† Copenhagen University Observatory, Øster Voldgade 3, DK-1350 Copenhagen K, Denmark

‡ Physics Department, University of Durham, South Road, Durham DH1 3LE, UK

§ Anglo-Australian Observatory, Epping Laboratory, PO Box 296, Epping, New South Wales 2121, Australia

OBSERVATIONS indicate that nearby supernova of type Ia have similar peak brightnesses, with a spread of less than 0.3 mag (ref. 1), so that they can potentially be used as 'standard candles' to estimate distances on a cosmological scale. As part of a long-term search for distant supernovae, we have identified and studied an event that occurred in a faint member of the distant galaxy cluster AC118, at a redshift of $z = 0.31$. Extensive photometry and some spectroscopy of the event strongly supports the hypothesis that we have detected a type Ia supernova whose time-dilated light curve matches that of present-day supernovae of this class. We discuss the precision to which its maximum brightness can be ascertained, and indicate the implications that such deep supernovae searches may have for observational cosmology.

Although supernovae are not as luminous as the brightest galaxies in clusters, they are events rather than objects and so should be less affected by the evolutionary and dynamical complications that have plagued determinations, by magnitude-redshift tests based on first-ranked cluster galaxies, of the deceleration parameter q_0 . If a sufficient number of supernovae

could be found and if they revealed a closely distributed (tight) Hubble diagram, precise photometry of a sufficiently deep sample could provide an interesting constraint on q_0 . The effect of a change in q_0 from 0.1 to 0.5 is only 0.13 mag at $z = 0.3$, rising to 0.22 mag at $z = 0.5$, so many accurately measured supernovae would be required. Our distant-supernova search programme has been described previously^{2,3}. Our recent estimate of the frequency of occurrence of type Ia supernovae³ lies at the lower end of the range determined in nearby galaxies⁴⁻⁶. Furthermore, even at maximum light such type Ia supernovae would be fainter than $V \approx 21.5$ mag, and thus any search strategy needs to reliably detect an absolute change in a galaxy's flux equivalent to $V \approx 23$ mag.

Using the 1.5-m Danish telescope at La Silla, Chile, we have monitored ~ 60 clusters in the redshift interval $0.2 < z < 0.5$ over a period of two years. One-hour CCD exposures in good conditions were taken during most months in the period August–April and were immediately compared with suitable frames taken at earlier epochs, by forming difference frames (after smoothing to the same seeing and scaling to the same object intensity)⁷. We have already discussed^{3,7} the discovery of a probable type II supernova at $z = 0.28$. That particular event was very faint but demonstrated our ability to find genuine events at $V \approx 23.6$ mag, a limit more than adequate for detecting type Ia supernovae to $z \approx 0.5$. Here we report the first detection of a type Ia supernova (SN1988U)⁸, the most distant so far discovered. Photometry of this object allows us to comment on the feasibility of estimating q_0 from a reasonable sample of such events.

The new event was identified in a faint galaxy in the field of the rich cluster AC118. This cluster was identified as part of the southern Abell catalogue⁹ and has since been extensively studied spectroscopically^{10,11}, although no previous spectrum exists of the galaxy in which this supernova occurred; the cluster redshift is 0.307. The event was found with the Danish 1.5-m telescope on 8 and 9 August 1988 by comparing a V CCD frame with one taken in good conditions during 1986 (Fig. 1). Observations at the 4.2-m William Herschel Telescope (WHT) on 9 August 1988 confirmed both the photometric detection and offsets measured at La Silla. Furthermore, the excess light has the same full width at half maximum (FWHM) as that for other stellar objects in the field. Subsequently, the cluster was observed several times on both the WHT and the 1.5-m Danish telescope when conditions and instrumentation permitted; a complete photometric record is given in Table 1.

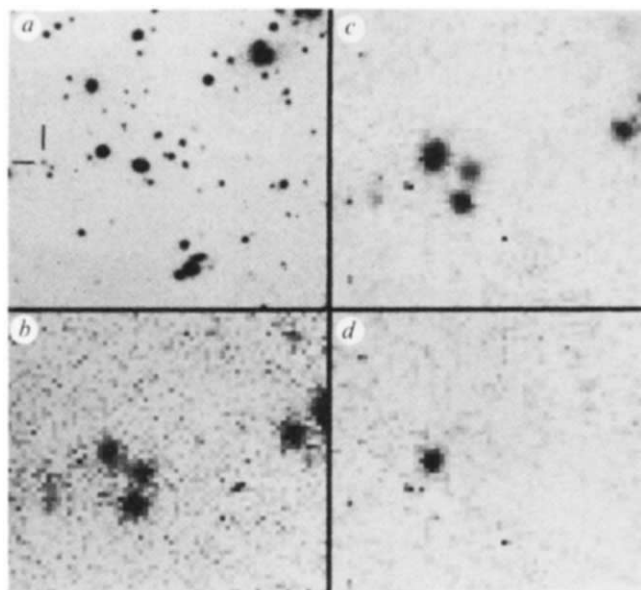


FIG. 1 Detection of the candidate supernova in the distant cluster AC118. *a*, 1-hour V exposure on the Danish 1.5-m telescope on 30 August 1986. *b*, 30×30 -arcsec enlargement of *a* around the galaxy containing the event. *c*, Same enlargement of a 1-hour V exposure taken on 8 and 9 August 1988. *d*, Subtracted image ($c - b$) after scaling and allowing for slightly different seeing on the two exposures. Analysis of the difference frame shows that the excess light is offset 0.5 arcsec east and 0.7 arcsec south of the galaxy nucleus. The galaxy position at right ascension $\alpha = 00^{\text{h}} 11^{\text{m}} 56.69^{\text{s}}$, declination $\delta = -30^{\circ} 41' 45.3''$ (1950.0). The galaxy has a redshift of $z = 0.31$ and $V = 22.35 \pm 0.03$ mag; the event was detected with $V = 22.05 \pm 0.05$ mag.

TABLE 1 Photometric record

Julian date (minus JD 2447373.5)	Aperture (m)	Seeing (arcsec)	Supernova V (mag)	ΔV
9.28	1.5	1.7	22.05	0.05
10.21	1.5	1.6	22.18	0.05
11.20	4.2	1.1	22.30	0.09
11.28	1.5	2.1	22.29	0.07
12.20	4.2	1.5	22.43	0.09
13.16	4.2	1.3	22.32	0.08
14.20	4.2	1.2	22.38	0.08
15.35	1.5	1.8	22.67	0.08
16.34	1.5	1.6	22.74	0.10
36.08	4.2	1.3	24.02	0.43
37.23	1.5	1.3	29.30	0.25
38.24	1.5	1.0	24.20	0.31
67.20	1.5	1.5	>24.4	—
<i>R</i>				ΔR
11.20	4.2	1.2	21.94	0.07
12.20	4.2	1.6	22.17	0.08
37.32	1.5	2.3	>24.1	—*
70.12	1.5	1.1	>24.4	—

* Estimated from frame taken in standard Gunn r filter²¹.

A spectrum of the event was secured with a 4,000-s integration using the Anglo-Australian Telescope (AAT) Faint Object Red Spectrograph at $\sim 20\text{-}\text{\AA}$ resolution on 19 August 1988 during time allocated to P. Hewett and S. Warren. The reduced spectrum was cross-correlated with that of the nearby supernova SN1981B (Type Ia), 17 days after maximum¹². We found a redshift $z=0.310$. The similarity (taking the contamination of the host galaxy into account) is striking, especially the characteristic Si II 6,150- $\text{\AA}$ absorption feature found generally in type Ia supernovae (Fig. 2). This strongly suggests that the event in AC118 is a type Ia supernova at the redshift of the cluster. A post-event spectrum of the galaxy was obtained with a 4,500-s integration using the AAT Low Dispersion Survey Spectrograph at $\sim 10\text{-}\text{\AA}$ resolution on 15 October 1988. A weak feature at wavelength $\lambda=5,200\text{ \AA}$ is consistent with the 4,000- $\text{\AA}$ break expected in a blue ($V-R=0.5$) galaxy at the cluster redshift.

To follow the supernova light curve, dark and bias CCD images were subtracted, flat-fielded and analysed using two independent techniques. First, a stellar photometric sequence within the CCD frames was established with respect to primary standards^{13,14} on a photometric night, and 5.2-arcsec-aperture photometry of the host galaxy was performed. The supernova brightness was derived by subtracting the value for the quiescent galaxy and photometric errors estimated, including photon noise and sky-subtraction contributions. As no R frame of AC118 was obtained before August 1988, the R magnitude of the quiescent galaxy was taken from the 8 October observation, on the assumption that the supernova makes an insignificant contribution at that date, as indicated by the V data. The second method used the technique described by Hansen *et al.*³, whereby each CCD frame is compared with a previous one that matches best in terms of seeing. Three-arcsec-aperture photometry of the difference frames, when corrected to total magnitudes using stellar objects in the original frames, shows excellent agreement with the first method above. However, the photometric errors are slightly better in the second method, and we have therefore adopted these values in Table 1.

We calculate the supernova light curve in terms of rest-frame B magnitudes, making proper allowance for the different filter bandwidths and adopting a $(B-V)$ -time relation appropriate for type Ia supernovae¹². This light curve agrees closely with that inferred from K -correcting the V observations using detailed spectra of SN1981B (ref. 15) and introduces a transformation error of ≤ 0.2 mag over the period concerned, an uncertainty less than the noise errors of Table 1. Figure 3 shows the rest-frame B light curve together with a template light-curve derived from nearby type Ia supernovae¹⁶ shifted to fit the

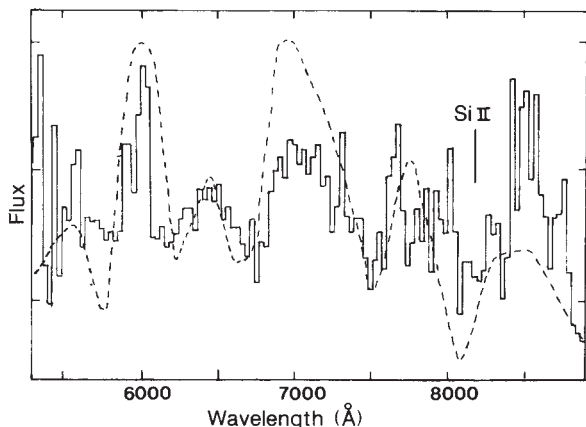


FIG. 2 Spectrum (solid line) of the galaxy in the cluster AC118, containing the supernova obtained by Hewett and Warren on 19 August 1988. The dashed line shows a spectrum of the nearby type Ia supernova 1981B 17 days after maximum redshifted to $z=0.31$. The redshifted 6,150- $\text{\AA}$ feature is visible at 8,060 $\text{\AA}$.

observed curve. Clearly the supernova has been discovered after maximum light. A least-squares linear fit to the decline defined by the first ten observations gives β (the decline rate of the light curve between the maximum and the inflection point) $= 7.1 \text{ mag} \pm 0.2 (100 \text{ days})^{-1}$, which lies in the low tail of the distribution of β values for nearby type Ia supernovae¹⁷⁻¹⁹, indicating that, without a correction for time dilation, the AC118 supernova would be highly atypical. Including time-dilation effects, however, we find $\beta=9.3$, locating the supernova near the middle of the β distribution. Thus the decline in brightness observed is consistent with a normal type Ia supernova and supports a time dilation that is proportional to $(1+z)$, which is consistent with a relativistic Friedmann cosmology.

Clearly the precision with which the peak brightness of the AC118 supernova can be established is hampered by its detection at post-maximum. As the β value is close to the normal value, we have used the standard light curve¹⁶ to obtain ΔB , the magnitude difference between maximum light and the inflection point on a type Ia supernova light curve; we find $\Delta B=2.8$. Although correlations between ΔB and β have been found by means of photographic photometry¹⁹, any correction to ΔB for the actual β value observed is less than 0.1 mag. On the assumption that the inflection point is at least as faint as the points at the observation time $t_{\text{obs}} \approx 37$ days on the light curve (Fig. 3), we find $B_{\text{max}} \geq 21.5 \pm 0.3$, and from the brightest observation we find a value $B_{\text{max}} \leq 22.4 \pm 0.2$. AC118 is included in the SA140 field, in which the galactic reddening has been estimated as $E(B-V)=0.055 \pm 0.029$ (ref. 20); we thus incorporate in our calculations an extinction $A_V=0.17 (\pm 0.09)$. If the absolute maximum brightness is $M_B = -18.18 \pm 0.13$ (ref. 12, and taking a Hubble constant $H_0=100 \text{ km s}^{-1} \text{ Mpc}^{-1}$), then the true distance modulus to AC118 has a range $39.5 \leq (m-M) \leq 40.4$ (ignoring the small uncertainties in the two estimates of B_{max} above), which is consistent with that expected for a reasonable value of q_0 (40.10–39.97 for $q_0=0.1\text{--}0.5$).

Although the AC118 supernova has behaved in all respects like a typical nearby supernova, we cannot yet judge whether it is a representative standard candle for our purposes, and hence a formal result for q_0 , however uncertain, is premature. The implied bounds, $-0.6 \leq q_0 \leq +2.5$ ($-0.9 \leq q_0 \leq +1.5$ with no extinction correction), are, however, sensible, considering the fact that maximum light was not observed directly. Increased statistics would therefore allow some interesting constraints to

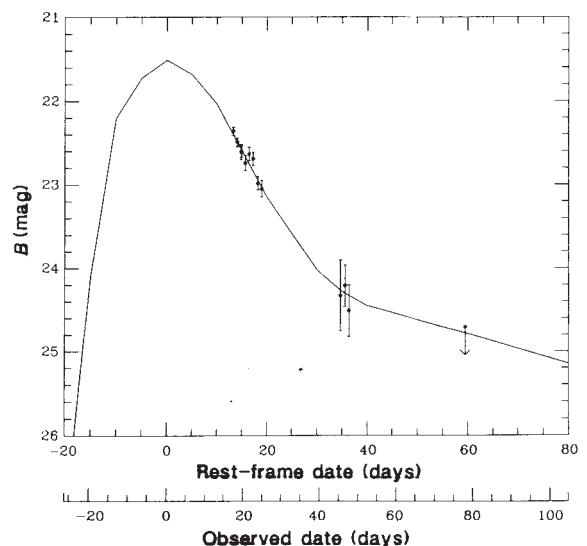


FIG. 3 Rest-frame B light curve of the AC118 supernova (points), uncorrected for extinction, compared with a standard light curve for a type Ia supernova (solid line)¹⁶. $t=0$ is chosen to coincide with the estimated epoch of maximum light.

be obtained as well as providing a measure of the degree of homogeneity of the sample. In particular, for supernovae discovered before maximum light, $B_{\max}$ should be determined to better precision. Calculations show that $\sim 40\%$ of supernovae found in our cluster sample (with an average redshift $\bar{z} \approx 0.3$) should be in this category. Even for post-maximum discoveries such as AC118, however, an accurate determination of the inflection point should be possible using the Hubble Space Telescope.

We have demonstrated that it is possible to extend the study of supernovae to cosmologically interesting distances ($z \leq 0.5$). By extending our survey to a cluster sample with $\bar{z} \approx 0.5$ we will need to detect ~ 15 supernovae to determine q_0 with an accuracy of 0.1, in the best possible case (that is, if the type Ia supernovae are completely well behaved objects with a cosmic dispersion of ~ 0.2 mag). To avoid uncertainties from internal reddening, type Ia supernovae in red members are preferred. Adopting our revised rate of occurrence of type Ia supernovae in distant clusters^{3,5}, updated to include observations up to October 1988, this will require $\sim 1,500$ comparisons. $\square$

Received 17 March; accepted 1 May 1989.

1. Sandage, A. & Tammann, G. A. in *Supernovae as Distance Indicators* (ed. Bartel, N.) 1–13 (Springer, Berlin, 1985).
2. Hansen, L., Nørgaard-Nielsen, H. U. & Jørgensen, H. E. *ESO Messenger* **47**, 46–49 (1987).
3. Hansen, L., Nørgaard-Nielsen, H. U., Jørgensen, H. E., Ellis, R. S. & Couch, W. J. *Astr. Astrophys. Lett.* **211**, L9–L11 (1989).
4. Tammann, G. in *Supernovae: A Survey of Current Research* (eds Rees, M. J. & Stoneham, R. J.) 371–403 (Reidel, Dordrecht, 1982).
5. Van den Bergh, S., McClure, R. D. & Evans, R. *Astrophys. J.* **323**, 44–53 (1987).
6. Cappellaro, E. & Turatto, M. *Astr. Astrophys.* **190**, 10–16 (1988).
7. Hansen, L., Jørgensen, H. E., Nørgaard-Nielsen, H. U., Ellis, R. S. & Couch, W. J. *IAU Circ. No.* 4641 (1988).
8. Hansen, L., Jørgensen, H. E., Nørgaard-Nielsen, H. U., Ellis, R. S. & Couch, W. J. *IAU Circ. No.* 4656 (1988).
9. Abell, G. O., Corwin, H. G. & Olowin, R. P. *Astrophys. J. Suppl.* (in the press).
10. Couch, W. J. & Newell, E. B. *Astrophys. J. Suppl.* **56**, 143–192 (1984).
11. Couch, W. J. & Sharples, R. M. *Mon. Not. R. astr. Soc.* **229**, 423–456 (1987).
12. Branch, D. *et al. Astrophys. J.* **270**, 123–139 (1983).
13. Landolt, A. U. *Astr. J.* **88**, 439–460 (1983).
14. Christian, C. A. *et al. Publ. astr. Soc. Pacif.* **97**, 363–372 (1985).
15. Leibundgut, B. thesis, University of Basel (1989).
16. Cadonau, R., Sandage, A. & Tammann, G. A. in *Supernovae as Distance Indicators* (ed. Bartel, N.) 151–165 (Springer, Berlin, 1985).
17. Pskovskii, Y. P. *Sov. Astr.* **21**, 675–682 (1977).
18. Branch, D. *Astrophys. J.* **258**, 35–40 (1982).
19. Pskovskii, Y. P. *Sov. Astr.* **28**, 658–664 (1984).
20. Knude, J. K. *Astrophys. Lett.* **18**, 115–118 (1977).
21. Thuan, T. X. & Gunn, J. E. *Publ. astr. Soc. Pacif.* **88**, 543–547 (1976).

ACKNOWLEDGEMENTS. We thank many observers for their cooperation in obtaining CCD observations of AC118 at short notice, including J. Teuber, D. Carter, J. Allington-Smith and T. Shanks. We are especially grateful to P. Hewett and S. Warren for securing the vital AAT spectrum. We thank Dr B. Leibundgut for providing us with his new supernova K-corrections.

Denitrification in the Antarctic stratosphere

R. J. Salawitch, G. P. Gobbi*, S. C. Wofsy & M. B. McElroy

Department of Earth and Planetary Sciences and Division of Applied Sciences, Harvard University, 29 Oxford Street, Cambridge, Massachusetts 02138, USA

RAPID loss of ozone over Antarctica in spring requires that the abundance of gaseous nitric acid be very low. Nitric acid is removed from the gas phase in the lower stratosphere at temperatures below about 195 K through the formation of crystalline nitric acid trihydrate^{1–4}, and below 188 K in association with ice crystals^{5–9}. Precipitation of particulate nitric acid has been assumed to occur in association with large ice crystals, requiring significant removal of H₂O and temperatures well below the frost point. However, stratospheric clouds exhibit a bimodal size distribution in the

Antarctic atmosphere, with most of the nitrate concentrated in particles with radii $\geq 1 \mu\text{m}$ (refs 10, 11). Here we argue that the bimodal size distribution sets the stage for efficient denitrification, with nitrate particles either falling on their own or serving as nuclei for the condensation of ice. Denitrification can therefore occur without significant dehydration: it is unnecessary for temperatures to drop significantly below the frost point.

The Airborne Antarctic Ozone Experiment (AAOE) conducted from Punta Arenas, Chile, in August and September 1987 provided unambiguous evidence that the removal of ozone over Antarctica in spring is effected mainly by reactions involving elevated levels of chlorine and bromine radicals¹². High concentrations of radicals imply an exceptionally low abundance of gaseous HNO₃. Otherwise, photolysis of HNO₃ provides a source of NO₂, allowing conversion of the reactive species ClO and BrO to the relatively unreactive ClONO₂ and BrONO₂.

Spectroscopic observations from McMurdo Station (78°S) in 1986^{13,14} gave the first quantitative evidence for loss of HNO₃ from the Antarctic stratosphere. *In situ* measurements by Fahey *et al.*^{15,16} in 1987 demonstrated loss of odd nitrogen (denitrification) and loss of H₂O (dehydration) in the Antarctic stratosphere. The mixing ratio of odd nitrogen (NO_x), normally in excess of 10 parts per 10⁹ by volume (p.p.b.v.) in unperturbed air, reached values as low as ~ 1 p.p.b.v. at an altitude of 18 km. Abundances of H₂O, ~ 4.5 p.p.m.v. in early winter, fell to values as low as 1.5 p.p.m.v. in early spring.

The condensation sequence in early winter initially involves the hydration of pre-existing sulphuric acid particles¹⁷. The sulphate aerosols may freeze at ≤ 220 K (ref. 17). Nitric acid, the most abundant component of NO_x, is expected to condense as crystalline nitric acid trihydrate (NAT) at temperatures ≤ 195 K at an atmospheric pressure level of 50 mbar (assuming abundances of 10 p.p.b.v. HNO₃ and 4.5 p.p.m.v. H₂O)^{1–4}. Particles of NAT with radii between 0.5 and 2 μm are believed to form the tenuous Type 1 polar stratospheric clouds (PSCs), with extinction coefficients generally $\sim 0.001 \text{ km}^{-1}$ (refs 1, 18). Water condenses at ~ 188 K, forming the larger crystals (3–8 μm) that make up Type 2 PSCs, with extinction coefficients often

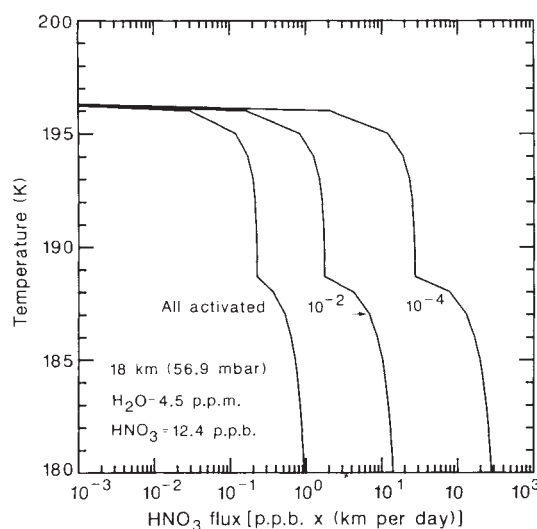


FIG. 1 The flux of nitric acid in sedimenting particles is shown as a function of temperature for various numbers of activated nuclei, expressed as a fraction of the initial condensation nuclei (assumed density 10 cm^{-3}). The air parcel was assumed to be at an altitude of 18 km and to maintain thermodynamic equilibrium with its environment at all times. Condensation of NAT begins at 196.3 K; the same particles activated during the growth of NAT were assumed to be activated when water ice forms (188.7 K). The analysis focuses on vapour pressures and particle densities in clouds after condensation, when vapour and solid phases will be approximately at equilibrium. Therefore the lower temperatures ($\Delta T \approx 1 \text{ K}$) required to activate background aerosols are not explicitly considered.

* Permanent address: Istituto di Fisica dell'Atmosfera CNR, CP 27-00044 Frascati, Italy.

$>0.01 \text{ km}^{-1}$. Measurements of particulate composition¹⁶ and lidar (laser radar) observations¹⁸ appear to confirm the formation of particles composed primarily of NAT in Type 1 PSCs, with ice the dominant component for Type 2 PSCs.

The condensation of nitric acid trihydrate or ice on a nucleus of different composition should be limited initially by the rate of nucleation^{19,20}. Nucleation requires a saturation ratio (ratio of ambient vapour pressure to equilibrium vapour pressure over the condensed phase) exceeding unity by an amount determined by the free-energy barrier that must be surmounted for nucleation. The largest aerosols seem to nucleate first^{10,11}, reflecting the lower free-energy barrier for larger condensation nuclei^{19,20}. Growth proceeds faster for larger nucleated particles because of the elevation of vapour pressure over smaller particles (the Kelvin effect)^{18–20}.

If cooling rates are sufficiently rapid, the saturation ratio can attain high values, thereby allowing nearly all of the condensation nuclei to grow. The size distribution would be unimodal, with condensate deposited more or less uniformly on all the nuclei. If cooling rates are slow, the first particles to nucleate can attain a surface area sufficient to take up the excess vapour. The saturation ratio may level off or decline, even as temperature continues to drop. Smaller aerosols may remain unactivated^{18–20} or may nucleate after much of the ambient vapour has been deposited on the largest particles. A bimodal distribution of particle sizes can be produced, with most of the condensed material associated with a small number of relatively large particles. A bimodal distribution can result also if the free-energy barrier to nucleation varies considerably among the condensation nuclei.

Bimodal distributions have been observed by Rosen *et al.*²¹ for ice crystals in Type 2 PSCs, and in tropospheric cirrus clouds forming at temperatures and pressures similar to those in PSCs⁸. Turco *et al.*⁶ have argued that condensation of NAT should involve essentially all available condensation nuclei, as a consequence of the strong temperature dependence of HNO_3 vapour pressure and the small size of stratospheric sulphate aerosols. Poole and McCormick¹⁸ argued for a bimodal distribution for NAT particles; however, they assumed very slow cooling rates (0.5 K per day) and did not have access to correct thermody-

namic data for NAT. The nucleation and growth of NAT on stratospheric sulphate aerosols is not well understood, and these discussions are necessarily preliminary. According to Turco *et al.*⁶, condensation of ice would involve a smaller fraction of the nuclei than condensation of NAT, and large ice crystals would contain only a small amount of the available nitric acid. Precipitation of ice particles would therefore represent an inefficient sink for HNO_3 .

Turco *et al.*⁶ addressed this issue by postulating that denitrification could occur in Type 2 PSCs forming in air masses subject to rapid cooling ($\sim 5 \text{ K per day}$). They suggested that the concentration of gaseous HNO_3 could remain significantly higher than the equilibrium vapour pressure during growth of Type 1 PSCs, allowing HNO_3 to be efficiently incorporated into growing ice crystals. It is important, however, that cooling rates are not too large: otherwise condensation of ice would involve a large fraction of the available condensation nuclei and the size of resulting particles would be limited to radii of a few micrometres, and sedimentation velocities would be correspondingly modest. If there is no significant disequilibrium of HNO_3 during the formation of Type 2 PSCs, dehydration would be expected to occur, accompanied by little or no denitrification.

Here we propose an alternative view, that denitrification occurs primarily as a consequence of differential nucleation and growth during formation of both NAT and ice particles. Observations by Hofmann *et al.*^{10,11} show that size distributions for Type 1 PSCs are often bimodal over Antarctica in spring, with most of the NAT mass deposited on $0.1\text{--}1\%$ of the pre-existing aerosol particles. Particle size distributions^{10,11} suggest condensation of NAT exclusively on the largest sulphate aerosols, producing a small number of NAT particles with radii between 1 and $2 \mu\text{m}$. We assume that mechanisms governing nucleation and growth in winter are similar to those in spring: the size distributions for NAT particles should consequently be bimodal in winter as well as spring. A bimodal distribution of NAT particles sets the stage for efficient denitrification in winter. Large particles of NAT can provide preferred sites for the deposition of ice (a point also made by Poole and McCormick¹⁸), or they can fall on their own. Removal of HNO_3 should be essentially complete after the first ice crystals precipitate. Denitrification can therefore be accomplished with little or no dehydration.

Consider for illustrative purposes an air parcel circulating in the polar vortex, undergoing adiabatic expansion (cooling) and compression (heating), at an altitude of $\sim 18 \text{ km}$ with a mean pressure of 56.9 mbar ; ref. 22. NAT becomes stable at this level at 196.3 K , ice at $\sim 188.7 \text{ K}$. We adopt a typical size distribution for stratospheric condensation nuclei²³ with a total of $10 \text{ particles cm}^{-3}$. The largest aerosols are assumed to nucleate first. We consider values from 100% to 0.01% for the fraction that nucleate.

Figure 1 shows the computed flux for precipitation of HNO_3 as air undergoes a condensation sequence through Type 1 and Type 2 PSCs. If all the aerosols nucleate, the mean particle radius is $0.4 \mu\text{m}$ for condensation of $10 \text{ p.p.b.v. HNO}_3$, increasing to $1.0 \mu\text{m}$ with additional condensation of $1 \text{ p.p.m.v. H}_2\text{O}$. Settling velocities are small. Loss of HNO_3 in the NAT phase is trivial, and, even in an ice cloud, $<1 \text{ p.p.b.v. of HNO}_3$ would be removed from a 1-km-thick layer in a day.

If 1% of the particles are activated, particles of NAT in the Type 1 PSC would have a density of 0.1 cm^{-3} and a mode radius near $1.7 \mu\text{m}$. Measurable vertical redistribution of HNO_3 could occur even without condensation of ice; $\sim 15\%$ of the available HNO_3 could be removed from a 1-km layer in a day. Ice particles in a consequent Type 2 PSC would have a mode radius $>4 \mu\text{m}$, reflecting condensation of $0.5 \text{ p.p.m.v. H}_2\text{O}$, or $5 \mu\text{m}$ for condensation of 1 p.p.m.v. Nitrate would be removed efficiently once the ice cloud formed: in one day, the large particles (containing most of the HNO_3) would fall 750 m if the mode radius was $5 \mu\text{m}$, or 500 m if the mode radius was $4 \mu\text{m}$. If the

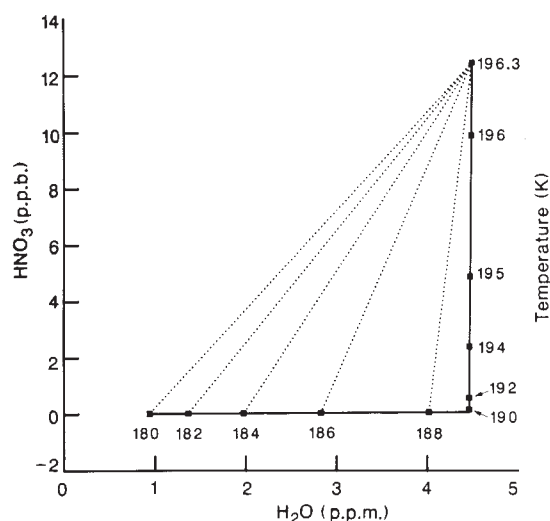


FIG. 2 Nitric acid/water relationships. The solid line shows the relationship between gas-phase HNO_3 and H_2O , as a function of temperature, for an air parcel at 18 km containing initially $12.4 \text{ p.p.b.v. of HNO}_3$ and $4.5 \text{ p.p.m.v. of H}_2\text{O}$. Equilibrium temperatures are indicated by squares along the solid curve⁴. The dotted lines represent relationships between HNO_3 and H_2O that result from mixing of denitrified and dehydrated air with unmodified air having the initial composition.

fraction of particles activated was as small as 10^{-3} or 10^{-4} , particles of NAT would have radii larger than 2.5 or 8 μm , respectively, allowing significant fluxes of nitrate over short timescales before the formation of ice clouds.

An illustrative model for the composition of the vortex in the winter of 1987 may be constructed using temperatures reported for the South Pole by Komhyr *et al.*²⁴. NAT would be stable at 18 km after 15 May, ice would be stable after 7 June. Temperatures fell below 185 K on 1 July, and remained below 188 K until September. The results in Fig. 1 indicate precipitation of NAT, before the formation of ice, if NAT is deposited on <0.1% of the available nuclei; ~ 10 p.p.b.v. of HNO_3 would be removed from a 7-km layer by 7 June. If NAT was deposited on 1% of the available nuclei, denitrification would be negligible until 7 June when the frost point was reached. The precipitation flux would exceed 5 p.p.b.v. of HNO_3 from a 1-km-thick layer per day by 15 June. A layer several kilometres deep would be denitrified by 15 June, with the loss of only 0.5 p.p.m.v. of water. Significant dehydration would occur after 15 June.

The sequence of denitrification and subsequent dehydration depends primarily on the particle size distribution for NAT. Details of the temperature history are unimportant, provided that a bimodal size distribution is produced for NAT. The prevalence of bimodal size distributions over McMurdo Station^{10,11} suggests that denitrification should have preceded dehydration over extensive regions of the Antarctic stratosphere in 1987.

Admixture of denitrified or dehydrated air with unprocessed air may obscure the sequence of condensation and precipitation, but a distinctive relationship between concentrations of HNO_3 and H_2O should be preserved. This is illustrated schematically in Fig. 2 (see also ref. 15). Concentrations of HNO_3 and H_2O would cluster on the right-hand side of Fig. 2 for air that has denitrified recently. Concentrations would cluster along the bottom of the triangular region for air that has denitrified and then dehydrated. Levels of HNO_3 and H_2O falling on the dotted lines in Fig. 2 would correspond to admixtures of denitrified-dehydrated air with unmodified air subsequent to the time when precipitation last occurred. Few points should plot outside the triangle, that is, we would expect that cases with high HNO_3 and low H_2O would be relatively rare. The observed levels of HNO_3 and H_2O provide information on the coldest temperature experienced by an air parcel over its history.

Figure 2 provides a framework for the analysis of HNO_3 and H_2O in the polar stratosphere, and allows a test of the hypothesis that denitrification can occur in advance of or in the absence of significant dehydration. The analysis above implies that the low temperatures needed to dehydrate the atmosphere may not be required for denitrification: halogen-catalysed destruction of ozone as observed over Antarctica can occur in the absence of appreciable dehydration. $\square$

Received 21 December 1988; accepted 27 April 1989.

1. Toon, O. B., Hamill, P., Turco, R. P. & Pinto, J. *Geophys. Res. Lett.* **13**, 1284–1287 (1986).
2. Crutzen, P. J. & Arnold, F. *Nature* **324**, 651–654 (1987).
3. McElroy, M. B., Salawitch, R. J. & Wofsy, S. C. *Geophys. Res. Lett.* **13**, 1296–1299 (1986).
4. Hanson, D. & Mauersberger, K. *Geophys. Res. Lett.* **15**, 855–858 (1988).
5. Toon, O. B., Turco, R. P., Jordan, J., Goodman, J. & Ferry, G. J. *Geophys. Res.* (in the press).
6. Turco, R. P., Toon, O. B. & Hamill, P. J. *Geophys. Res.* (in the press).
7. Ramaswamy, V. *Geophys. Res. Lett.* **15**, 863–866 (1988).
8. Heymsfield, A. J. *J. Atmos. Sci.* **43**, 851–855 (1986).
9. Stanford, J. L. *J. Atmos. Sci.* **30**, 1431–1436 (1973).
10. Hofmann, D. J., Rosen, J. M., Harder, J. W. & Hereford, J. V. *J. Geophys. Res.* (in the press).
11. Hofmann, D. J., Rosen, J. M. & Harder, J. W. *NASA Conf. Publ. No. 10014*, 57–59 (1988).
12. Anderson, J. G. *et al. J. Geophys. Res.* (in the press).
13. Farmer, C. B., Toon, G. C., Schaper, P. W., Blavier, J.-F. & Lowes, L. L. *Nature* **329**, 126–130 (1987).
14. McElroy, M. B., Salawitch, R. J. & Wofsy, S. C. *Planet. Space Sci.* **36**, 73–87 (1988).
15. Fahey, D. W. *et al. J. Geophys. Res.* (in the press).
16. Fahey, D. W. *et al. J. Geophys. Res.* (in the press).
17. Steele, H. M., Hamill, P., McCormick, M. P. & Swisler, T. J. *J. Atmos. Sci.* **40**, 2055–2067 (1983).
18. Poole, L. R. & McCormick, M. P. *J. Geophys. Res.* **93**, 8423–8430 (1988).
19. Pruppacher, H. R. & Klett, J. D. *Microphysics of Clouds and Precipitation* (Reidel, Hingham, 1980).
20. Fletcher, N. H. *The Physics of Rainclouds* (Cambridge University Press, 1962).
21. Rosen, J. M., Hofmann, D. J., Carpenter, J. R., Harder, J. W. & Oltmanns, S. J. *Geophys. Res. Lett.* **15**, 859–862 (1988).

22. Chan, K. R., Scott, S. G., Bui, T. P., Bowen, S. W. & Day, J. J. *Geophys. Res.* (in the press).
23. Pinnick, R. G., Rosen, J. M. & Hofmann, D. J. *J. Atmos. Sci.* **33**, 304–314 (1976).
24. Komhyr, W. D., Franchois, P. R., Kuester, S. E., Reitelbach, P. J. & Fanning, M. L. *NOAA Data Rep. ERL ARL-15* (NOAA Environmental Research Labs, Boulder, 1988).

ACKNOWLEDGEMENTS. We thank D. Fahey for helpful discussions. This work was supported by the NSF and by NASA. G. P. G. was supported in part by a NATO Senior Fellowship and by the Italian Antarctic Program while visiting Harvard University.

Non-equilibrium molecular evaporation of carboxylic acid dimers

M. Faubel & Th. Kisters

Max-Planck-Institut für Strömungsforschung, Bunsenstrasse 10, D-3400 Göttingen, FRG

IN molecular liquids that may form dimers, for example by hydrogen-bonding, direct evaporation of dimers is possible when these form stable molecules. A gradual change from monomer to dimer evaporation would be expected at pairwise molecular bond energies of intermediate strength—that is, ~ 0.01 – 1 eV, a range of values that is typical of hydrogen-bond strengths in liquids. Here we use molecular-beam techniques to study the evaporation of large fractions of acetic and formic acid dimers from the pure liquids and from solutions. We find that their velocity distributions are non-maxwellian, and that the average dimer energies exceed the temperature of the liquid surface by 100–200 K. These unexpected effects may be qualitatively explained by a crude model in which surface-tension forces act upon non-wetting, twin-hydrogen-bond dimer inclusions within an extended, monomeric hydrogen-bonded 'sheet' at the liquid surface.

For common hydrogen-bonded liquids, including water and acetic acid, the equilibrium vapour pressure is several torrs and corresponds to a mean free path λ of the order of 10 μm . The molecular evaporation process may therefore be studied under collision-free, vacuum-evaporation conditions by restricting the liquid surface to such small physical dimensions¹. For molecular-beam sources, the hydrodynamic supersonic gas expansion that develops at high pressures and large beam-source diameters, $D \gg \lambda$ (ref. 2), transforms into a maxwellian, collisionless free molecular flow (characteristic of that in a Knudsen cell) when D is reduced to dimensions comparable to λ . A controlled steady-state liquid surface with a stationary, axially decreasing temperature profile is obtained by injecting a fast-flowing thin jet of liquid into a vacuum³. The disappearance of vapour-phase collisions for very thin liquid jets is experimentally verified by time-of-flight (TOF) measurements of the velocity distribution of a collimated molecular-beam sample of the vapour that emerges from a small region of the liquid jet. This experimental configuration is illustrated in Fig. 1a; Fig. 1b gives TOF measurements that show the transition from collision-dominated vapour flow to free evaporation for thin, 10- μm jets of pure water.

For quantitative analysis, we fit to the TOF spectra a standard floating maxwellian velocity distribution function of supersonic molecular flow²:

$$f(v) dv = Cv^2 \exp \left[\frac{m(v-u)^2}{2kT_{\parallel}} \right] \quad (1)$$

Here $T_{\parallel}$ is the local gas temperature, u is the average flow velocity, m is the molecular mass, v is the molecular velocity, and C is a constant. The 'speed ratio' $S = (mu^2/2kT_{\parallel})^{1/2}$ provides a convenient criterion for the extent of molecular collisions in the expanding gas. For $S = 0$ the velocity distribution is the ordinary maxwellian distribution for collision-free molecular flow. The molecular-beam source enthalpy, and thus the local

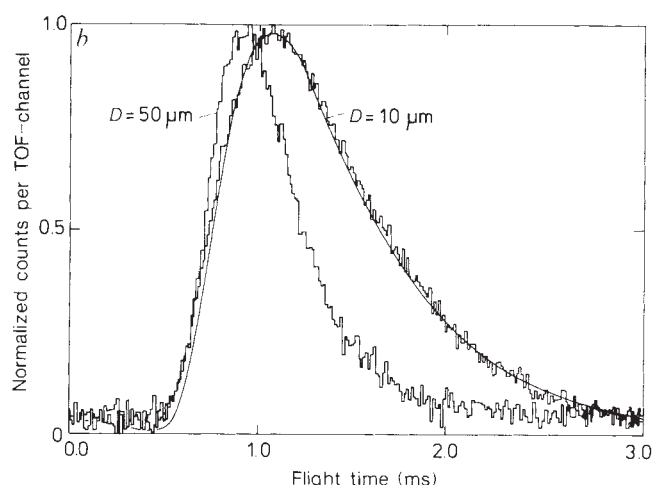
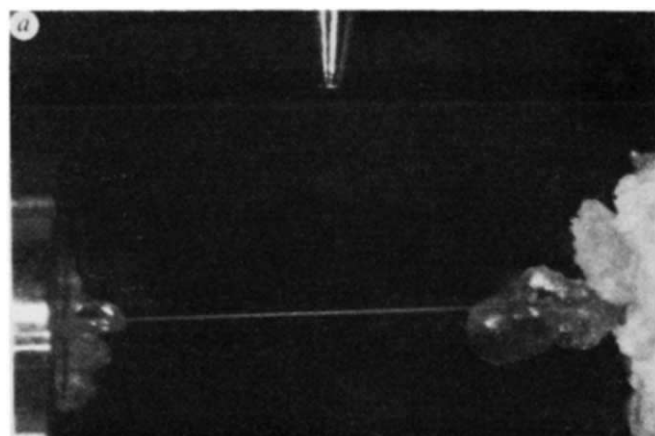


FIG. 1 From the free surface of a very thin liquid jet, molecules can evaporate into a vacuum without vapour-phase collisions when the jet diameter is smaller than the free molecular path ($\sim 10\text{ }\mu\text{m}$ for water or acetic acid). *a*, The thin liquid jet is injected at room temperature by a high-pressure nozzle (100 bar) (left). Ice deposits cover the entrance of the cryo-pumped beam dump. Vapour molecules are sampled perpendicular to the beam for molecular-beam velocity analysis. The 0.7-mm-aperture conical skimmer defines the entrance to the time-of-flight spectrometer. *b*, Representative water-vapour time-of-flight spectra, showing a narrow, supersonic velocity distribution for a 50- μm liquid jet and a much wider distribution for a 10- μm jet (flight path is 80 cm). These spectra confirm the expected transition from nozzle beam to free molecular flow when the ratio between jet diameter and mean free path approaches the Knudsen limit. The smooth line shows the theoretical Maxwellian distribution.

surface temperature of the evaporating liquid jet, is $T_s = [0.4 \times S^2 + 1] T_{\text{fl}}$, assuming that internal molecular energy is not relaxed.

We have used this molecular-beam vapour-sampling method to investigate possible direct evaporation of dimers or larger clusters from a liquid surface. For H_2O , no dimers were detected within the sensitivity limits of 10^{-4} of the monomer intensity, although $(\text{H}_2\text{O})_2$ dimers with 120-meV binding energy⁴ form readily in conventional molecular-beam sources^{5,6}. Carboxylic acids are known to form exceptionally stable dimers with a 0.6-eV twin hydrogen bond between the two COOH groups⁷. Neutron scattering indicates that acetic acid contains planar dimers, although the solid has an infinite-chain crystal structure⁸. Known ionization and fragmentation patterns of acetic acid dimers and trimers⁹⁻¹¹ facilitated our quantitative study of dimeric evaporation. 10- and 20- μm -diameter liquid jets of acetic acid may be studied readily by the technique illustrated in Fig. 1*a*. The TOF spectra of the nascent acetic acid vapour show the evaporation of both monomers and dimers (Fig. 2*a, b*). With the mass resolution $m/\Delta m = 50$ of the molecular-beam detector used here, the CH_3COOH monomer parent mass of 60 was not fully separated from the intense dimer-fragment peak at mass 61. Therefore the monomer TOF spectrum had to be taken at the 'clean' fragment mass 45, where less than 5% of the peak intensity originates from dimer fragments. The dimer spectrum was measured on ion fragment mass 105. The mass-spectrum intensity ratios indicate a dimer concentration of $\sim 15\%$, using the fragmentation values of ref. 11. The absence of a signal at the trimer fragment masses of 103 and 121 implies that no trimers evaporate.

Analysis of the CH_3COOH monomer TOF spectrum in Fig. 2*a* gives a best-fit speed ratio of $S^2 = 0.2$ and a beam-enthalpy-derived surface temperature of $T_s = 252\text{ K}$. However, a dramatically larger best-fit speed ratio of $S^2 = 2.2$ is found for the dimer TOF distribution in Fig. 2*b*. Moreover, this non-Maxwellian dimer velocity distribution shows a best-fit translational enthalpy that corresponds to a source temperature of 365 K, which is $\sim 100\text{ K}$ above the surface temperature of the liquid jet.

This unexpected result is confirmed by further measurements in which the observed surface cross-section, the surface temperature, the local gas density, the jet diameter and the liquid composition were varied. These measurements are summarized in Table 1. A particularly clear experimental demonstration of the absence of gas-phase collisional effects in the formation of this non-Maxwellian velocity distribution of the evaporating dimers is provided by the TOF spectra for a liquid mixture of 20% acetic acid in 80% water. In this case, both the water monomer and the acetic acid monomer show Maxwellian distributions with near-zero speed ratios (Fig. 2*c*). The surface temperatures of $T_s = 281\text{ K}$ for the water and $T_s = 275\text{ K}$ for the CH_3COOH monomer are equal within experimental accuracy. As the average molecular velocities for H_2O and CH_3COOH

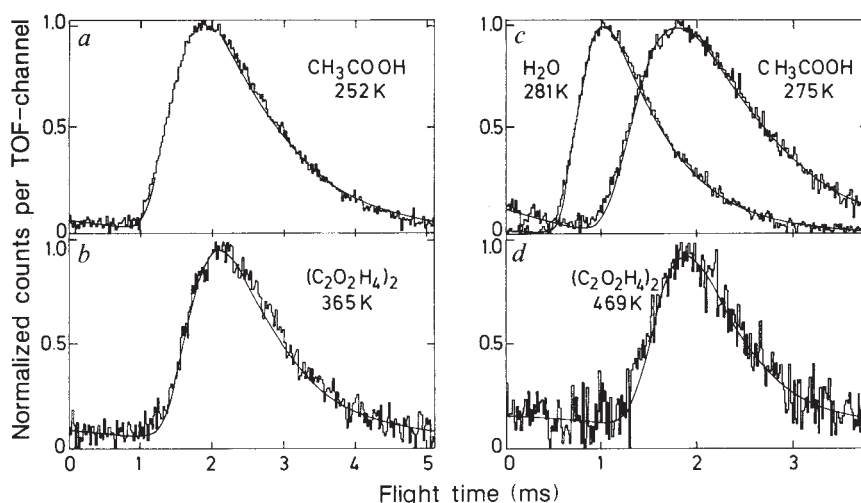
TABLE 1 Source temperature and speed ratios of monomer and dimer TOF distributions

Liquid	Beam diameter (μm)	Solvent		Monomers		Dimers	
		S^2	$T_s(\text{K})$	S^2	$T_s(\text{K})$	S^2	$T_s(\text{K})$
CH_3COOH (99.8%)	10	—	—	0.3	258	2.2	369
	10	—	—	0.4	245	3.5	399
	20	—	—	0.2	265	2.8	376
$\text{CH}_3\text{COOH} + \text{H}_2\text{O}^*$	10	0.4	251	0.4	271	7.2	429
	20	0.1	279	0.4	272	6.0	472
$\text{CH}_3\text{COOH} + \text{C}_2\text{Cl}_4^*$	10	0.3	270	—	—	[0.0	350]
	20	0.1	310	0.3	255	[0.0	408]
HCOOH	10	—	—	0.5	250	2.9	411

Square brackets indicate data with high noise levels, caused by C_2Cl_4 fragments.

* 20% CH_3COOH : 80% solvent. About 30% (by weight) of the evaporating acid is found to be in the form of dimers.

FIG. 2 Liquid acetic acid evaporates both as monomer ($\sim 70\%$) and as dimer molecules ($\sim 30\%$ by weight). The monomer velocity distribution is maxwellian. The TOF spectra of dimers can be fitted only by large-speed-ratio floating maxwellians. Dimer source temperatures (T_s) always exceed the monomer-derived temperatures by 100–200 K. *a* and *b*, TOF spectra for simultaneously evaporating monomers ($S^2=0.2$, $T_s=252$ K) and dimers ($S^2=2.2$, $T_s=365$ K) of 99.8% pure acetic acid. Jet diameter is 20 μm , injection pressure is 80 bar, observation point is 10 mm downstream, skimmer distance is 1 mm. *c* and *d*, 10- μm jet of a mixture of 20% acetic acid in 80% H_2O . The monomer temperatures (*c*) agree within experimental accuracy ($S^2=0.1$, $T_s=281$ K for H_2O and $S^2=0.3$, $T_s=275$ K for CH_3COOH); for the co-evaporating dimers (*d*), $S^2=5.2$ and $T_s=469$ K.



differ by almost a factor of two, any vapour collisions would manifest themselves in a hydrodynamic acceleration and increased temperature of the heavier monomers². As before, however, the distribution of simultaneously evaporating $(\text{CH}_3\text{COOH})_2$ dimers (Fig. 2*d*) is clearly non-maxwellian. Note that in this case both the speed ratio and the translational enthalpy excess of the dimers are considerably larger than those observed for pure acetic acid.

A complete theoretical explanation for this phenomenon is not yet available. The observed dimer energy distribution can be formally explained by an activation-energy barrier for dimer evaporation. Before a molecule evaporates it is bound in a potential well with a depth equal to the molecular heat of evaporation, E_v . In the normal evaporation process the molecular potential energy rises abruptly across the liquid-to-vacuum boundary. Molecules with an equilibrium velocity distribution within the region of the potential well retain their maxwellian distribution and their original temperature when they evaporate across such a simple potential step^{12,13}. When, however, a potential barrier, with height E_a , is also present at the liquid boundary the evaporating molecules have first to overcome this 'activation' step by attaining an energy $E'_v = E_v + E_a$, thus being ejected with an energy E_a on the vacuum side of the boundary. This activation energy is added to the kinetic energy of all molecules and leads to a shifted maxwellian distribution (equation (1)) with $\frac{1}{2}mu^2 \approx E_a$.

On a microscopic, molecular level the observed dimer energy excess could originate from the 0.6-eV energy release in the dimer formation process, in a manner analogous to a recently reported reaction mechanism for the (non-equilibrium) desorption of H_2 molecules from solid metal surfaces^{14,15}. Inaccuracies in the available heats of evaporation and dimer binding-energy data, however, prevent estimates of such effects on a scale of

~ 100 K. In addition, such a mechanism would have to occur at a very high rate on a timescale comparable to that for the evaporation of a liquid monolayer (10^{-7} s). We therefore propose a different, essentially macroscopic model in which the dimer disequilibrium originates from surface-tension effects at the liquid/dimer interface. When cyclic dimers coexist with chain-like associations of monomers within the liquid⁸, the hydrophilic COOH groups will be 'saturated' in the centre of the dimers. The dimer molecules then appear as hydrophobic, non-wetting particles within the monomeric acetic acid solvent. The surface-tension forces between the ambient liquid monomers will act like a stretched 'trampoline', expelling the dimers in order to minimize the hydrogen-bond-linked surface of tension, as depicted in Fig. 3. The gain in surface energy when a dimer is expelled is roughly $E_{\text{th}} = \sigma\pi r^2$, where σ is the surface tension and the dimer is assumed initially to form a hemispherical surface deformation with radius r . The values of the surface tensions ($\sigma = 1.7$, 3.1 and 2.7 $\text{meV } \text{\AA}^{-2}$ for acetic acid, 20% acetic acid in water and formic acid¹⁶ respectively) are, in fact, roughly proportional to the observed dimer temperature excesses of 100 K, 200 K and 150 K for the respective liquids. A surface reduction corresponding to $\pi r^2 \approx 6 \text{\AA}^2$ gives the correct order of magnitude of the effect within this crude model.

It would be interesting to investigate these phenomena further by realistic molecular-dynamics simulation. It is possible that this observation of so large a fraction of dimers with 30–50% higher-than-average thermodynamic energy may have implications for macroscopic properties such as solvent/solvate properties, or may affect the thermodynamic description of the liquid/vapour interface of these carboxylic acids and perhaps of hydrogen-bonded liquids more generally. □

Received 27 February; accepted 10 May 1989.

- Knudsen, M. *Ann. Phys.* **28**, 75–130 (1909).
- Anderson, J. B. in *Molecular Beams and Low Density Gas Dynamics* (ed. Wegener, P. P.) (Dekker, New York, 1974).
- Faubel, M., Schlemmer, S. & Toennies, J. P. *Z. Phys.* **D10**, 269–277 (1988).
- Reimer, J. R. & Watts, R. O. *Chem. Phys.* **85**, 83–112 (1984).
- Vernon, M. F. *et al. J. chem. Phys.* **77**, 47–57 (1982).
- Dyke, T. R., Mack, K. M. & Muentner, J. S. *J. chem. Phys.* **66**, 498–510 (1977).
- Chao, J. & Zwolinski, B. J. *J. phys. Chem. Ref. Data* **7**, 363–377 (1978).
- Bertagnoli, H. *Chem. Phys. Lett.* **93**, 287–292 (1982).
- Sievert, R., Cadez, I., van Doren, J. & Castleman, A. W. *J. phys. Chem.* **88**, 4503–4505 (1984).
- Mori, Y., Kitagawa, T., Yamamoto, T., Yamada, K. & Nagahara, S. *Bull. chem. Soc. Jap.* **53**, 3492–3495 (1980).
- Mori, Y. & Kitagawa, T. in *Atomic Collision Research in Japan* **13**, 138–141 (Society for Atomic Collision Research, Riken, Saitama 351–01, 1987).
- Boltzmann, L. *Wien. Ber.* **63**, 397–420 (1871).
- Richardson, O. W. *Phil. Mag.* **28**, 633–647 (1914).
- Comsa, G., David, R. & Rendulic, D. *Phys. Rev. Lett.* **38**, 775–778 (1977).
- Comsa, G. & David, R. *Surf. Sci. Rep.* **5**, 145–198 (1985).
- CRC Handbook of Chemistry and Physics* (ed. West, R. C.) (CRC Press Inc., Florida, 1977).

ACKNOWLEDGEMENTS. We thank S. Schlemmer, J. P. Toennies and V. Vogel for discussions, and J. Pick for technical assistance.

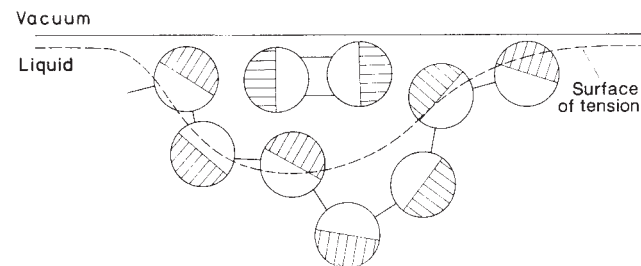


FIG. 3 Schematic representation of the crude surface-tension model used to explain the uniform energy excess of dimer molecules. The dimers represent hydrophobic inclusions within the chain-like structure of hydrogen-bonded monomers. To minimize the surface of tension of the monomeric liquid, dimers are expelled from the surface with an energy proportional to the surface energy of a bubble with a radius equal to that of the dimer.

Lidar detection of leads in Arctic sea ice

R. C. Schnell*, R. G. Barry*, M. W. Miles*,
E. L. Andreas†, L. F. Radke‡, C. A. Brock‡,
M. P. McCormick§ & J. L. Moore||

* Cooperative Institute for Research in Environmental Sciences (CIRES),
University of Colorado, Boulder, Colorado 80309, USA

† US Army CRREL, Hanover, New Hampshire 03755, USA

‡ University of Washington, Seattle, Washington 98195, USA

§ NASA, Langley Research Center, Hampton, Virginia 23665, USA

|| Wyle Corporation, Hampton, Virginia 23666, USA

REMOTE sensing using an airborne infrared lidar¹ has shown an unexpected capability to detect open leads (linear openings) in Arctic sea ice and their associated meteorology in winter. Here we show that vertical profiles of backscattered radiation demonstrate strong returns from hydrometeor plumes originating from leads having a surface water temperature near -1.8°C . Recently refrozen leads are also distinguishable by the lidar backscatter from adjacent thicker, older sea ice. Wide leads release enough energy to create buoyant plumes which penetrate the Arctic boundary layer inversion, transporting heat and moisture into the troposphere. These results show that the role of the Arctic as a global heat sink may need to be re-evaluated, and that lead plumes have a significant effect on the radiation budget.

A hydrometeor plume rising from an open lead and a plume streamer from another lead are shown in Fig. 1 from an Arctic flight by the National Aeronautics and Space Administration (NASA) Electra², $\sim 83\text{--}87^{\circ}\text{N}$, 70°W , on 27 January 1984. The rising plume (left side of Fig. 1) has an apparent ice source ~ 300 km north of Ellesmere Island, and reaches an altitude of nearly 4 km. The streamer to the south probably comes from a wide lead off Ellesmere Island, which is not in Fig. 1 but was observed on the return flight. This streamer extended downwind (northward) for ~ 250 km and reached a height of 4 km. The layer of aerosol observed at 3 km (16:37 to about 16:50 GMT) is thought to be Arctic haze air pollution. As the backscatter signals in the plumes are strongest near the surface and decrease upward, these features are not consistent with precipitating clouds. Because of these unexpected observations, a further study of the phenomenon was undertaken.

On 15 April 1986, as part of the Arctic Gas and Aerosol Sampling Program (AGASP)³, a research aircraft operated by the Cloud and Aerosol Research Group of the University of Washington, equipped with cloud-physics instrumentation⁴ and a $1.06\text{-}\mu\text{m}$ downward-pointing lidar⁵, obtained cloud-microphysical data over open leads near Thule, Greenland. The lead-induced clouds were clearly associated with open-water regions. A transect through the cloud tops showed the presence of ice crystals $< 50\text{ }\mu\text{m}$ in diameter, at a temperature of -25°C . The bulk of the condensation below the aircraft had the visual appearance of cloud (stratocumulus) containing liquid water, although the aircraft did not penetrate it. Condensation immediately above the surface of open leads is in the form of water droplets $\sim 10\text{ }\mu\text{m}$ in diameter⁶. These may exist as supercooled droplets at temperatures as low as -40°C , depending on the presence of freezing nuclei or ice multiplication mechanisms⁷.

The lidar system flown on the aircraft was also able to identify newly frozen leads. A flight made on 18 April 1986 off Baffin Island, Canada, (Fig. 2) shows both refrozen leads and a hydrometeor plume from an open lead. These lidar data were processed in a backscatter ratio colour spectrum designed to highlight ice features. In Fig. 2, the surface appears as a sharp boundary near the bottom edge; below the surface is a region of colours fading down the backscatter scale. Large areas with multiple leads and single leads with dimensions greater than the horizontal resolution of the lidar (40 m) lack the gradual

colour gradient immediately 'beneath' the surface. This gradient results because when the laser beam from the lidar strikes the highly reflective snow surface, a strong return pulse saturates the receiver, which takes many nanoseconds to recover, thus producing a false signal below the surface. The less reflective surface of open and thinly covered leads produces no such saturation and thus no false signal. The 'below surface' signal should be interpreted as a measure of reflectivity, not of ice thickness. The location and presence of the open and refrozen leads was confirmed by photography and visual observations.

Leads are of considerable interest to oceanographers and atmospheric scientists. Much of the heat lost from the Arctic in winter is converted into ice in the leads; the resulting salt rejection and brine drainage from the ice helps to maintain the Arctic Ocean halocline and may drive secondary circulation patterns. Leads result from ice deformation, primarily under wind forcing, and have typical widths of 10–100 m. Analyses of aircraft photographs, Landsat imagery and submarine sonar data on under-ice draft show that the size distribution of leads in Arctic sea ice follows a negative power law such that along a 100-km track we could expect to encounter 30 leads that are 50 m wide but only 3 that are 500 m wide⁸. In winter, in the absence of ice deformation or movement of frazil ice by the wind, leads rapidly refreeze, forming new ice within a few hours^{9,10}. Although open water and areas of new ice represent only a few per cent of the ice-covered Arctic Ocean in winter, they have a major role in turbulent heat transfer to the atmosphere^{11,12}. Refrozen leads with ice 0.2–0.8 m thick may account for more than half of the turbulent energy transfer from the ocean to the atmosphere in winter in the central Arctic^{9–13}.

Previous field measurements^{6,12,14,15} of the turbulent scalar fluxes over leads suggest that the heat and moisture transferred to the atmosphere remains within the atmospheric boundary layer because of the strong low-level temperature inversion which dominates the lower Arctic troposphere in winter. Our observations show, however, that highly energetic plumes can penetrate to an altitude of at least 4 km.

The upper-air sounding from Alert, Ellesmere Island, for 12:00 GMT 27 January 1984—five hours before the observations in Fig. 1—shows an inversion of 8°C between the surface and a pressure level of 850 mbar with an equivalent potential temperature (θ_e) at 4 km of 0°C . As θ_e is conservative for adiabatic processes, to penetrate the inversion and reach 4 km, a plume originating from a lead must be warmed and moistened at the surface to a θ_e of 0°C . We have developed a simple model to see whether such drastic modification of the surface-level air is possible.

The heat and moisture escaping from leads generally remain within an internal boundary layer (tens to hundreds of metres deep), which grows in height in the along-wind direction (x)

TABLE 1 Model results

Fetch (m)	$\bar{T}$ ($^{\circ}\text{C}$)	θ_e ($^{\circ}\text{C}$)	$10^4 C$ (kg m^{-3})	$10^9 N$ (m^{-3})	H_s (W m^{-2})	H_l (W m^{-2})
50	-23.7	-21.6	8.3	1.6	218	52
100	-22.1	-19.5	9.1	1.7	214	51
500	-17.4	-13.7	10.4	2.0	204	49
1,000	-15.3	-11.0	10.4	2.0	197	48
5,000	-9.1	-3.1	8.2	1.6	182	45
10,000	-7.0	-0.4	6.4	1.2	169	43
20,000	-4.2	3.1	3.2	0.4	160	41

$\bar{T}$ is near-surface air temperature; θ_e , equivalent potential temperature; C , condensate concentration; N , number of condensate droplets, assuming each has a diameter of $10\text{ }\mu\text{m}$; H_s and H_l , surface-averaged fluxes of sensible heat (enthalpy) and latent heat, respectively. The water temperature is -1.8°C ; the upwind air temperature, -35.8°C ; and 10-m wind speed, 2.5 m s^{-1} . The near-surface absolute humidity is always the saturation value at $\bar{T}$.

FIG. 1 Backscatter signal from a 1.06- μm downward-viewing lidar flown on the NASA Electra north of Ellesmere Island on 27 January 1984. White indicates maximum backscatter signal, decreasing linearly to black, which is indicative of scattering from gas molecules only. Shown are an apparent rising plume of hydrometeors (left) and a plume of aerosol (centre and right) that had a surface origin at a near-shore lead 250 km to the south off Alert. The aerosol layer at 3 km is thought to be Arctic haze air pollutants.

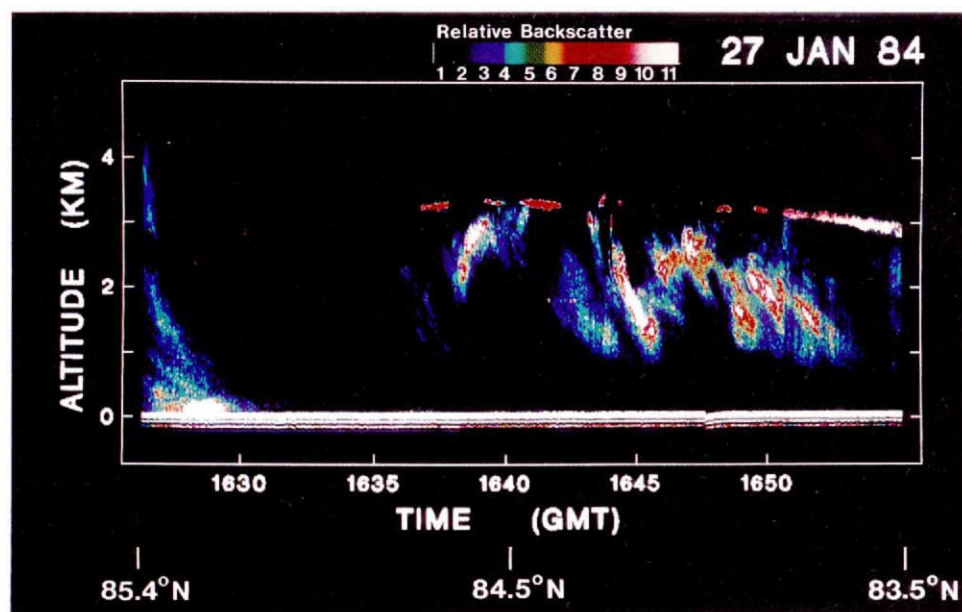
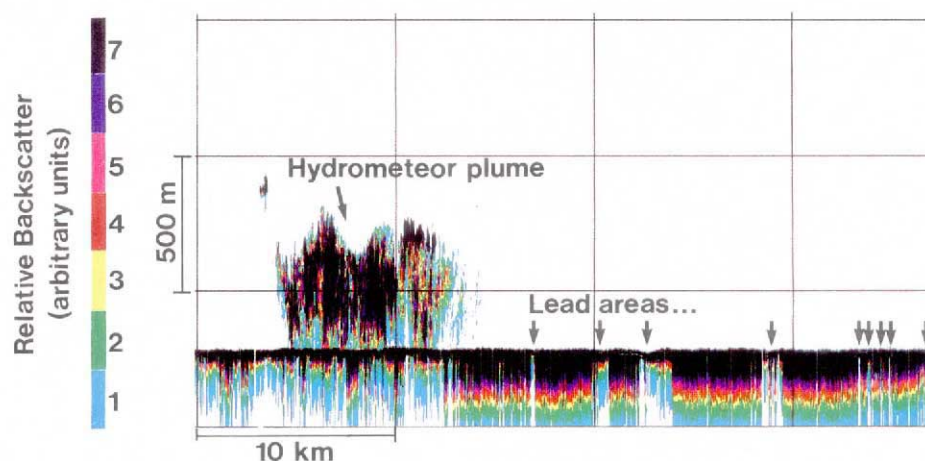


FIG. 2 False-colour imagery of the relative backscatter signal from the lidar operated on the aircraft. Colours indicate backscattering intensity from the lead-produced hydrometeor plume and the ice surfaces. The signal is corrected and scaled to eliminate returns from gas molecules and aerosol particles. Lower backscatter beneath the ice surface indicates refrozen leads.



approximately as $\delta(x) = 0.22 x^{0.8}$ (ref. 16). Using this idea, Andreas *et al.*¹⁴ developed conservation equations for the sensible (H_s) and latent (H_L) heat lost by leads. Andreas and Murphy¹⁷ recently showed how to estimate these surface-averaged heat fluxes by bulk aerodynamic formulas. Putting these results together, we have derived a prediction equation for the average air temperature in the lowest $\alpha\delta$ of the internal boundary layer ($\alpha = 0.05$),

$$\bar{T} = T_i + (XC_H/\alpha\delta)(T_s - T_i)$$

where T_i is the average upwind air temperature, T_s is the water surface temperature, X is the fetch (distance travelled by the wind) across the lead, and C_H is the bulk aerodynamic transfer coefficient for sensible heat¹⁶. A similar equation can be derived for the average water vapour content ($\bar{Q}$) of the layer.

Table 1 shows values of $\bar{T}$, θ_e (including freezing), the condensate concentration C , and the average sensible and latent heat fluxes predicted by this model. We computed C and the number of condensate droplets N by converting all the water vapour $\bar{Q}$ in excess of the saturation limit into droplets of radius 5 μm (ref. 6).

According to Table 1, a lead must be ~ 10 km wide to warm and moisten the air flowing over it to a θ_e of 0°C . Realistically, however, this is the narrowest lead that could produce a plume that reaches 4 km for conditions in the Alert sounding. A plume

will not rise adiabatically; through entrainment it will lose heat and moisture to the colder, drier ambient air. Therefore, to reach 4 km, the plume must have a surface θ_e value in excess of 0°C ; a more sophisticated plume model, however, is necessary to predict how much in excess of 0°C θ_e should be. Although ten similar major leads, estimated to be 5–10 km wide, are visible on Defense Meteorological Satellite Program infrared (2.7-km resolution) imagery along the flight path for 27 January 1984, only three produced plumes; the others had evidently recently refrozen.

It is not surprising that such extensive plumes have not been recognized before. They are not necessarily visible to the naked eye. Also, previous research has focused on relatively narrow open-water areas^{6,9–14}. Our calculations suggest that small leads are unable to warm and moisten the air enough to allow for vertical transport through a strong inversion.

Our observations are significant for three reasons. First, airborne lidar provides a capability for locating narrow leads that are smaller than the resolution of current satellite sensors. Recently refrozen leads are as easily detectable by lidar as the plumes from open leads. This capability is important for operational forecasting and for fuller understanding of ice dynamics. Second, the current view that turbulent fluxes from leads and polynyas affect only the boundary layer needs modifying. If heat and moisture from leads can regularly reach the

mid-troposphere, the role of the Arctic as a global heat sink may need re-evaluating and climate models will require more realistic parametrizations of surface-atmosphere fluxes. Third, our results suggest that lead plumes may affect the radiation budget. Using observed ice-crystal spectra, Curry *et al.*¹⁸ calculated a significant heating effect at the surface, due to the downward infrared radiation from the ice crystals. Table 1 shows that even narrow leads produce particle concentrations larger than those assumed by Curry *et al.*¹⁸. Thus, plumes will affect the radiation budget in their immediate vicinity and in large areas downwind.

Airborne and satellite-borne lidar provide the potential for determining the impact of leads on the energy and moisture fluxes from the polar ocean to the atmosphere. Further research will focus on developing a climatology of lead occurrence¹⁹ and of plume frequency and distribution in the Arctic from lidar profiles and satellite data, identifying the phase of the plume moisture with lidar polarization measurements, and on developing parametrizations to relate lidar backscatter to plume moisture content. □

Received 27 February; accepted 4 May 1989.

1. Utse, E. E., Morley, B. M. & Nielsen, N. B. *Appl. Opt.* **21**, 460–463 (1982).
2. Kent, G. S., Poole, L. R. & McCormick, M. P. *J. Atmos. Sci.* **43**, 2149–2166 (1986).
3. Schnell, R. C. *Arctic Res.* **2**, 39–41, 1988.
4. Hobbs, P. V. & Rango, A. L. *J. Atmos. Sci.* **42**, 2523–2549 (1985).
5. Radke, L. F., Brock, C. A., Lyons, J. H., Hobbs, P. V. & Schnell, R. C. *J. Atmos. Chem.* (in the press).
6. Andreas, E. L., Williams, R. M. & Paulson, C. A. *Q. J. R. Met. Soc.* **107**, 437–460 (1981).
7. Mason, B. J. *The Physics of Clouds* 2nd edn, 171–172 (Clarendon, Oxford, 1971).
8. Wadhams, P. *Phil. Trans. R. Soc. Lond. A* **302**, 45–85 (1981).
9. Badgley, F. I. *Proc. Symp. Arctic Heat Budget and Atmospheric Circulation* 267–277 (Rand Corporation, Santa Monica, 1966).
10. Andreas, E. L. *Mon. Weath. Rev.* **108**, 2057–2063 (1980).
11. Maykut, G. A. *J. Geophys. Res.* **87**, 7971–7984 (1982).
12. Makhtas, A. P. *The Heat Budget of Arctic Ice in the Winter*, (ed. Andreas, E. L.) (National Science Foundation, Division of Polar Programs, Washington, DC, in the press).
13. Ledley, T. S. *J. Geophys. Res.* **93**, 15919–15932 (1988).
14. Andreas, E. L., Paulson, C. A., Williams, R. M., Lindsay, R. W. & Businger, J. A. *Bound. Layer Met.* **17**, 57–91 (1979).
15. Smith, S. D., Anderson, R. J., den Hartog, G., Topham, D. R. & Perkin, R. G. *J. Geophys. Res.* **88**, 2900–2910 (1983).
16. Andreas, E. L. *CRREL Rep. 82-12* (US Army Cold Regions Research and Engineering Laboratory, Hanover, 1982).
17. Andreas, E. L. & Murphy, B. J. *phys. Oceanogr.* **16**, 1875–1883 (1986).
18. Curry, J. A., Radke, L. F., Brock, C. A. & Ebert, E. E. *Symp. Role of Clouds in Air Chemistry and Global Climate* 114–117 (American Meteorological Society, Boston, 1989).
19. Barry, R. G., Miles, M. W., Cianflone, R. C., Scharfen, G. & Schnell, R. C. *Ann. Glaciol.* **12**, 9–15 (1989).

ACKNOWLEDGEMENTS. This work was supported by the Office of Naval Research, the NSF's Division of Polar Programs, and NASA. The lidar used on the University of Washington aircraft was owned and operated by SRI International, Menlo Park, California, under contract to the AGASP project supported by NOAA.

The abrupt termination of the Younger Dryas climate event

W. Dansgaard*, J. W. C. White† & S. J. Johnsen‡

* Geophysical Institute, University of Copenhagen, Haraldsgade 6, DK-2200 Copenhagen N, Denmark

† Center for Geochronological Research, INSTAAR, University of Colorado, Boulder, Colorado 80309-0450, USA

‡ Science Institute, University of Iceland, Dunhaga 3, IS-107 Reykjavik, Iceland

PREVIOUS studies on two deep Greenland ice cores have shown that a long series of climate oscillations characterized the late Weichselian glaciation in the North Atlantic region¹, and that the last glacial cold period, the Younger Dryas, ended abruptly 10,700 years ago². Here we further focus on this epoch-defining event, and present detailed heavy-isotope and dust-concentration profiles which suggest that, in less than 20 years, the climate in the North Atlantic region turned into a milder and less stormy regime, as a consequence of a rapid retreat of the sea-ice cover. A warming of 7 °C in South Greenland was completed in about 50 years.

The disintegration of the North American and Eurasian ice sheets lasted approximately from 15,000 to 8,000 yr BP. During

this transition period, the well-known Bølling/Allerød—Younger Dryas (B/A-YD) climate oscillation occurred, the last of a long series of similar events that characterized the late Weichselian glaciation¹. The B/A-YD has been recorded much more distinctly in Europe and Greenland than in North America, indicating that it was closely connected to changing physical conditions in the North Atlantic Ocean, for example rapidly changing sea-ice cover³. The Younger Dryas was the last glacial cold spell, and it therefore seems reasonable to identify the termination of the Weichselian glaciation as the Younger Dryas—pre-Boreal transition, particularly because this was an abrupt and well defined event, which has been absolutely dated at 10,720 ± 150 yr BP, (8,770 ± 150 yr BC)².

In Greenland ice cores, the transition is characterized by a sudden shift in nearly all parameters studied so far; for example, concentrations of heavy stable isotopes⁴, chemical trace elements⁵, and acidity and continental dust⁶ all reflect radical changes in the Arctic environment. This is exemplified in the left-hand side of Fig. 1, where two previously published Greenland ice-core records (*b* and *c*) are compared with a European lake-core record (*a*).

The $\delta^{18}\text{O}$ profile in Fig. 1*a* was measured on calcite sediments along a core from Lake Gerzen, Switzerland⁷. The ^{14}C datings are given in conventional ^{14}C years. Comparison with the independent absolute dating at 10,720 yr BP (deduced by counting annual ice layers downwards from the surface) suggests that the conventional ^{14}C scale should be corrected by +720 ± 200 years at 10,000 ^{14}C yr, which is essentially in agreement with other estimates².

Figure 1*b* shows the $\delta^{18}\text{O}$ profile measured on ice in the 1,700–1,855-m depth interval of the deep ice core from Dye 3 in South Greenland, which was drilled under the American–Danish–Swiss joint effort, Greenland Ice Sheet Program, 1979–81⁸. The oxygen isotope profile is considered to reflect the mean annual surface temperature on the ice sheet. It spans most of the general warming from glacial to interglacial conditions, which was interrupted by the B/A-YD. The two completely independent records, Fig. 1*a* and *b*, agree even in detail, leaving no doubt that the B/A-YD oscillation followed similar paths in Europe and in Greenland.

Figure 1*c* shows the concentration of wind-blown continental dust in the ice, which increases with the storminess (frequency and intensity of storms), and with the extension and dryness of the source areas of the dust, for example the North American loess areas exposed to wind scouring. The dust concentration varies in antiphase with $\delta^{18}\text{O}$, which suggests dry and very stormy weather conditions in periods of full glacial severity, such as the Younger Dryas, but more humid and calm conditions in the mild interstadials, such as Bølling/Allerød (compare ref. 1 for data on previous interstadials).

Hence, the general atmospheric circulation in the Northern Hemisphere must have shifted dramatically, and certainly much faster than the extinction of the two large ice sheets mentioned above.

Here we present new isotopic evidence which indicates that aspects of this final transition must have been extremely fast, on the order of a couple of decades.

The termination of Younger Dryas at about 10,700 yr BP is illustrated in detail by the three curves *d–f* in Fig. 1, in which the depth scale, expanded nearly 40 times, spans now only the 1,784–1,788.2-m interval.

Figure 1*d* shows that the $\delta^{18}\text{O}$ level shifted by 5‰ within a 50-yr period. This timescale is derived from the fact that the *in situ* annual layer thicknesses just below and just above the 1.1 m of ice containing the δ -shift have been measured at 2 and 3 cm, respectively². With some reservations¹, the shift in the $\delta^{18}\text{O}$ value can be translated directly into a change in the surface temperature. The 5‰ increase in δ corresponds to a 7 °C warming⁹, that is, more than half of the total 12 °C Pleistocene-to-Holocene warming in South Greenland¹⁰. A 7 °C warming within

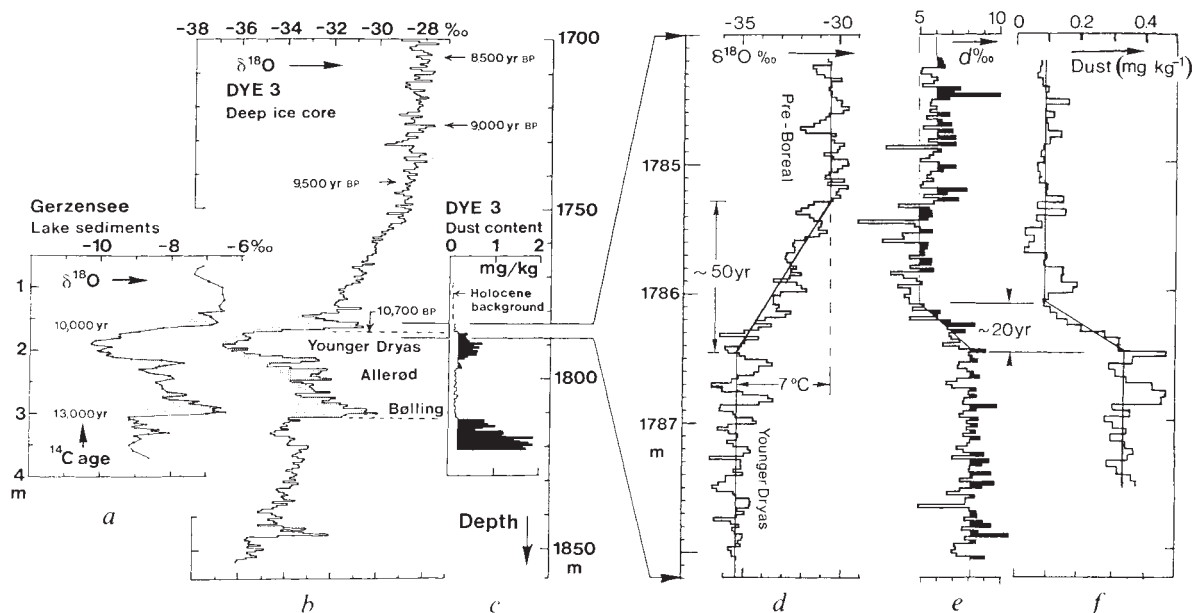


FIG. 1 *a*, Radiocarbon-dated $\delta^{18}\text{O}$ profile along a 4-m-long sediment core from Lake Gerzen in Switzerland. *b*, $\delta^{18}\text{O}$ profile along 150 m of the deep ice core from Dye 3 in South Greenland. The 1,700–1,850-m depth interval spans the entire Pleistocene-to-Holocene transition, including the Bølling/Allerød–Younger Dryas oscillation, which is revealed in essentially the same way as in the Swiss lake sediments (*a*). *c*, The concentration of

continental dust changed in anti-phase with $\delta^{18}\text{O}$ in Pleistocene ice. *d*, Detailed $\delta^{18}\text{O}$ record through the Younger Dryas–Pre-Boreal transition, during which the South Greenland temperature increased by 7°C in ~ 50 yr. *e* and *f*, Deuterium excess and dust concentration shifted to lower levels in less than 20 yr, probably because of retreating sea-ice cover, which left the North Atlantic Ocean with a period of calmer climate.

half a century is fast from a climatological point of view, but not unrealistic in view of the 2°C and 4°C warming that occurred over a decade in South and North Greenland, respectively, during this century (in the 1920s).

The deuterium and ^{18}O components, δD and $\delta^{18}\text{O}$, usually vary in parallel in the water cycle, δD changing 8 times faster than $\delta^{18}\text{O}$ in terms of ‰ deviation relative to standard mean ocean water (SMOW). However, kinetic effects during evaporation into undersaturated air, and condensation from supersaturated air, cause deviations from the simple relationship $\delta\text{D} = 8 \times \delta^{18}\text{O}$ (ref. 11). These deuterium excesses ($\delta\text{D} - 8\delta^{18}\text{O}$), denoted by d , are also influenced by the temperature of evaporation in the source areas of the moisture (d increases by $\sim 0.4\%$ for every $^\circ\text{C}$ warming of the surface sea water^{4,12}).

A deuterium excess as high as 8‰ in snow falling at high elevations on the Greenland ice sheet has recently been shown⁴ to be compatible only with high sea surface temperatures (18 – 20°C) in the main source areas of the moisture. Hence, in the lower part of Fig. 1*e*, the d level of 8‰ in ice from Younger Dryas suggests a dominating moisture source in the subtropical part of the Atlantic Ocean, which was then situated some 10° of latitude further south than now, according to CLIMAP 1976¹³.

Figure 1*e* further shows that, simultaneously with the beginning of the warming (compare Fig. 1*d*), d dropped to 5‰ within no more than 20 yr. A possible explanation is that when the North-Atlantic sea-ice margin, and thereby the polar front, retreated rapidly from its Younger Dryas position at 55° latitude³, vast areas of relatively cold surface sea water opened up as a temporary contributor of moisture. This would lead to low deuterium-excess in snow on the ice sheet, perhaps predominantly as a summertime phenomenon at first. Higher d -values were gradually restored over the next 30 yr, as the North Atlantic Ocean warmed up, with a relatively stable value of 6‰ established some 50 yr after the beginning of the end of the Younger Dryas, at the same time as the $\delta^{18}\text{O}$ of the snow reached a relatively stable value (Fig. 1*d* and *e*; it should be noted that both deuterium excess and $\delta^{18}\text{O}$ did not reach their full Holocene values of 8‰ and -27.5% until considerably later).

^{14}C dating of ocean sediment cores suggests that the retreat of the polar front from 35°N to 55°N took less than 400 yr (ref. 14), and indeed the duration of the transition appears from Fig. 1*e* to have been only 20 yr.

According to Fig. 1*f*, the dust concentration decreased in the same 20-yr period by a factor of about three, reaching values only twice the general Holocene background value of $0.05\text{ mg solid particles per kg ice}$. This may also be explained by the retreat of the North Atlantic sea-ice cover, which established a wide belt of open sea water of intermediate temperatures in between the tropical and polar water masses. In other words, the latitudinal temperature gradients, and thereby the pressure gradients, the storminess, the atmospheric turbidity and the deposition of continental dust on the ice sheet, decreased. Later on, vegetation covered the loess areas south of the retreating American ice sheet, thus reducing an important dust source. This is at least part of the explanation for the successive further reduction of the dust concentrations in the Greenland ice sheet.

In conclusion, Greenland ice cores reveal abrupt and radical changes in the North Atlantic region during the Younger Dryas–Pre-Boreal transition, including decreased storminess, a 50% increase in precipitation rate, a 7°C warming, and probably a temporary decrease of the evaporation temperature in the source area of moisture ultimately precipitated as snow at high elevations in the Arctic. □

Received 20 March; accepted 4 May 1989.

1. Dansgaard, W. *et al.* *M. Ewing Symp.* **5**, (eds Hansen, J. E. & Takahashi, T.) *Am. geophys. Un. geophys. Monogr.* **29**, 288–298 (1984).
2. Hammer, C. U., Clausen, H. B. & Tauber, H. *Radiocarbon* **28**, 284–291 (1986).
3. Ruddiman, W. F. & McIntyre, A. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **35**, 145–214 (1981).
4. Johnsen, S. J., Dansgaard, W. & White, J. *Tellus* **41B**(4) (in the press).
5. Herron, M. M. & Langway, C. C. in *Greenland Ice Core* (eds Langway, C. C., Oeschger, H. & Dansgaard, W.) *Am. geophys. Un. geophys. Monogr.* **33**, 77–84 (1985).
6. Hammer, C. U. *et al.* in *Greenland Ice Core* (eds Langway, C. C., Oeschger, H. & Dansgaard, W.) *Am. geophys. Un. geophys. Monogr.* **33**, 77–84 (1985).
7. Oeschger, H. *et al.* *M. Ewing Symp.* **5** (eds Hansen, J. E. & Takahashi, T.) *Am. geophys. Un. geophys. Monogr.* **29**, 299–306 (1984).
8. *Greenland Ice Core* (eds Langway, C. C., Oeschger, H. & Dansgaard, W.) *Am. geophys. Un. geophys. Monogr.* **33** (1985).
9. Dansgaard, W. & Oeschger, H. in *The Environmental Record in Glaciers and Ice Sheets* (eds Oeschger, H. & Langway, C. C.) *Dahlem Workshop* **8**, 287–317 (Wiley, New York, 1989).

10. Dahl-Jensen, D. & Johnsen, S. J. *Nature* **320**, 250–252 (1986).
11. Jouzel, J. & Merlivat, L. *J. geophys. Res.* **89**, 11749–11757 (1984).
12. White, J. W. C., Johnsen, S. J. & Dansgaard, W. *Ann. Glaciol.* **10**, 219–220 (1988).
13. CLIMAP Project Members *Science* **191**, 1131–1137 (1976).
14. Bard, E. *et al. Nature* **328**, 791–794 (1987).

ACKNOWLEDGEMENTS. This work was funded primarily by the Danish Natural Science Research Council, The Iceland Science Foundation and the US NSF, Division of Polar Programs. The authors are further indebted to Dr Jouzel for discussions and permission to let one of us (J.W.) do the deuterium measurements in his stable-isotope laboratory at Saclay.

High-precision mass-spectrometric uranium-series dating of cave deposits and implications for palaeoclimate studies

W.-X. Li*, J. Lundberg†, A. P. Dickin*, D. C. Ford†, H. P. Schwarcz*, R. McNutt* & D. Williams‡

* Department of Geology, McMaster University, Hamilton, Ontario L8S 4M1, Canada

† Department of Geography, McMaster University, Hamilton, Ontario L8S 4K1, Canada

‡ Trigon Computer Services, Melbourne, Florida 32901, USA

CALCITE deposits in caves, known as speleothem, can provide valuable palaeoenvironmental information^{1–4}. In particular, because speleothem are deposited only in air-filled caves, gaps in deposition in coastal caves record high-stands of sea level. Here we report the use of isotope-dilution mass spectrometry to date speleothem by the uranium-series method. The use of mass spectrometry in uranium-series dating was first applied to corals⁵, and has greatly improved the precision of this dating method. The speleothem dated here—a flowstone from 15 m below modern sea level in a Bahamian cave—records changes in sea level over the past 280,000 years. The dated hiatuses in deposition indicate high sea-level stands that are in general agreement with data from deep-sea oxygen isotope stratigraphy⁶ and other estimates for the timing of high-stands and glacial minima^{1–3,7}.

Corals record only the highest sea stands; speleothem complement the coral record in that they record a much larger fraction of the total range of sea level. Coral has the advantages of high U content (approximately 3 p.p.m.), easy recognition of alteration of the primary aragonite to calcite, and negligible incorporation of common (detrital) ²³⁰Th. Although generally having lower U content (<1 p.p.m.), speleothem is typically precipitated as coarsely crystalline calcite which inhibits recrystallization or chemical migration of isotopes. Most speleothem from the deep interiors of caves contain negligible amounts of common ²³⁰Th.

Clear calcite crystals were dissolved in 7 M HNO₃. A ²²⁹Th/²³⁶U tracer (calibrated against uraninite in secular equilibrium) was added. The solution was dried, redissolved in 7 M HNO₃ and loaded on a pre-cleaned and conditioned 12-ml nitrate-form anion-exchange column. Major-element ions were eluted with 7 M HNO₃. U and Th were eluted with H₂O and 1 M HBr and purified further through two sets of 0.25-ml columns. The final stage involved separation of Th (eluted with 6 M HCl) from U (eluted with 1 M HBr). The dried U fraction was loaded onto a tantalum-rhenium double-filament bead. The Th fraction was loaded onto the V-shaped side filament of a rhenium triple-filament bead, because this yielded the best Th⁺ ion emission. All filaments were of zone-refined metal. Filament beads were outgassed under vacuum before loading. Blanks showed no U or Th counts higher than the background level.

Isotope ratios were measured on a VG354 solid-source mass spectrometer with a 27-cm radius and 90° magnet sector. Uranium and thorium analyses used a Daly multiplier detector

in the analogue mode. Ions of all the isotopes of interest were measured sequentially. Counts were corrected for background (measured at mass 227.3) and for interpeak interference (measured at 0.5 mass units above and below each peak). Background noise was extremely low and tailing was negligible. A stable ion beam could be maintained for several hours with no significant fractionation: for Th runs of 2–3 h, means of each of the sets of 10 ratios deviated from the grand mean of the whole run by less than 0.2%. ²³⁰Th could be measured to ±0.13% (1σ) and ²³⁴U to ±0.05%. The errors were propagated through the standard dating equations⁸. The resulting 1σ uncertainty in age was often less than 1%.

Uranium was analysed in two stages: isotope masses 235, 236 and 238 over 60 cycles and then masses 234, 235 and 236 over

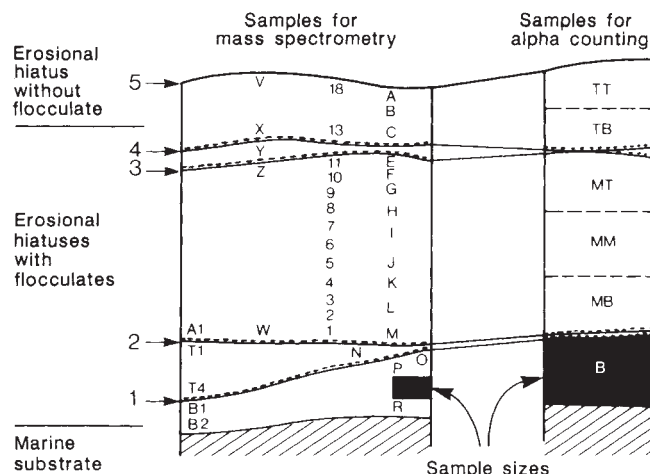


FIG. 1 Sampling diagram for Bahamas flowstone DWBAH. The dashed lines show the four erosional hiatuses with FeO(OH) flocculate. The uppermost erosional surface has no flocculate because sea level remains above it. Shaded areas compare sample sizes for the two dating methods. Sample size is a function of U content and age: for the DWBAH speleothem (with a U content of 0.15 p.p.m.), samples of 3–5 g sufficed for mass-spectrometric analysis, shown in black on the left, whereas alpha-spectrometric analysis, on the right, would require ~40 g of sample.

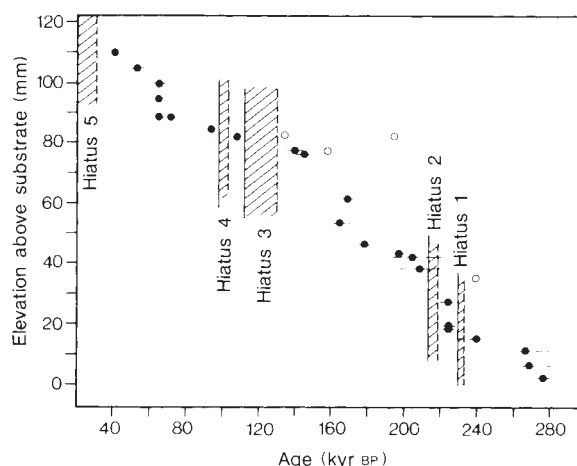


FIG. 2 Mass-spectrometric dates for Bahamas flowstone sample DWBAH from ~15 m in a submerged cave. Dates are shown by solid circles with 1σ error bars. Where no error bar is apparent it is smaller than the size of the symbol. Open circles indicate samples on hiatuses that show unexpectedly high ages as a result of U leaching. Vertical solid lines show the date of sea-level fall. Vertical dashed lines show the approximate date of sea-level rise. The cross-hatched areas show periods of non-deposition during high sea stands.

100 cycles. Each run was monitored and adjusted to keep the ^{234}U ion beam high and the Ta filament current at the optimum for U ionization, between 2.0 and 2.3 A. This single-spike, two-stage method for ^{234}U measurement differs from the double-spike, single-stage method of Edwards *et al.*⁵.

Thorium was normally run in one stage over 100 cycles (masses 229, 230 and 232) but if detrital Th content was high, interference from the high ^{232}Th necessitated a two-stage run (masses 229, 232 and then 229, 230). The centre-filament Re current was kept at ~ 5 A to give a Re^+ ion-beam current of 1.5×10^{-11} A. Each run was monitored to yield a stable and long-lived ^{230}Th beam with current $\sim 5 \times 10^{-17}$ A, while keeping the side-filament Re current between 2.2 and 2.6 A.

The measured ratios must be corrected for mass discrimination by the Daly detector and for mass fractionation during ionization at the filament. Ideally both are corrected by normalizing a measured isotopic ratio to an accurately known constant value. For the first stage of U analyses the naturally occurring, constant $^{235}\text{U}/^{238}\text{U}$ ratio of 0.0072526 was used. The corrected $^{235}\text{U}/^{236}\text{U}$ ratio then became the normalizing ratio for the second stage.

Normalization of Th was done as follows. Measurements of NBS U500 standard on the Faraday detector (which shows no mass discrimination) indicated that fractionation at the filament was only -0.007% per mass unit, whereas Daly measurements on NBS U500 showed that fractionation plus mass discrimination was up to 0.5% per mass unit. We assumed that Th behaves in a similar fashion to U, that fractionation is also negligible compared to mass discrimination at the detector, and that the measured ratios can be adequately corrected with the mass discrimination factor of $\sqrt{(m_1/m_2)}$.

Analyses of two standards demonstrated the precision of this method (Table 1a). We then dated the submerged flowstone

from the Bahamas (Table 1b), where the growth record helps to constrain eustatic changes in sea level. Previous alpha-spectrometric studies^{1-3,7} showed how dates on drowned speleothem can set limits on times of high sea stands on Bermuda and the Bahamas. The sample chosen was of high purity (low detrital Th); also, its low U content (0.15 p.p.m.) provided an extreme test of the method.

The sample (Fig. 1), from 15 m (± 0.75 m) below sea level in Lucayan Caverns, Grand Bahama Island, shows three prominent, and two lesser, growth layers. Each is terminated by an erosional hiatus. Upon each hiatus, except the final one, is a thin layer of goethite ($\text{FeO}(\text{OH})$) flocculate. This pattern is found to be the predominant flowstone sequence to depths of -20 m in caves of the Bahamas Banks.

The sequence of (1) flowstone deposition, (2) dissolution of its surface, (3) flocculate deposition and (4) resumption of flowstone deposition represents (1) calcite deposition in the vadose zone, (2) subaqueous solution, probably at the fresh/salt-water interface, (3) salt-water flocculation of iron compounds and (4) return to vadose conditions. The hiatuses are products of sea-level rise. The flowstone immediately above each goethite layer dates sea-level fall through -15 m or a little more (to allow for a shallow fresh-water lens); however, the date of each sea-level rise cannot be pinpointed so confidently because an unknown amount of flowstone is lost by dissolution during inundation and submergence of the cave.

Figure 2 shows the mass spectrometric dates and Table 1b illustrates their much greater precision and reproducibility compared with alpha counting. Except for samples on the hiatuses (which probably experienced U leaching during ensuing dissolution events), all the dates are in chronological order within the error bars. Bearing in mind tectonic subsidence of the Bahamas

TABLE 1

(a) Tests of standards					
NBS005a*	$^{234}\text{U}/^{238}\text{U}$	2σ			
NBS value	3.417×10^{-5}	0.070×10^{-5}			
McMaster value† (mean of 6)	3.416×10^{-5}	0.003×10^{-5}			
McMaster standard speleothem 76001‡					
Technique	N	Mean	Age (kyr) 1σ	Low age	High age
Alpha count	29	47.35	2.85	44.50	50.20
Mass spectrometry	12	47.58	1.17	46.40	48.75

(b) Comparison of alpha-counted ages and some mass-spectrometric ages for equivalent samples of the DWBAH speleothem							
Sample		Alpha-count dates		Mass-spectrometric dates			
		Age (kyr)	1σ error	Sample	Age	1σ error	
2DWBAH	BHB	>350		B2	274	+4.7	-4.7
	B	>350		P	241	+3.5	-3.5
	2B	>350					
Hiatuses 1 and 2							
	MB	170	+69 -42	A1	209	+11.0	-14.0
	MBR	307	+64 -119	1	197	+3.5	-3.5
	MM	203	+121 -55	6	168	+2.0	-2.0
	MT	139	+54 -39	10a	144	+5.2	-4.6
				10b	139	+4.8	-5.3
Hiatuses 3 and 4							
	TB	66	+11 -10	14	71	+2.0	-2.0
	TBR	60	+6 -6	15	64	+0.5	-0.5
	TT	76	+8 -7	18	39	+0.7	-0.7

* Uranium isotopic standard NBS005a (National Bureau of Standards, USA). Our measured ratios were well within the error of the NBS figures.

† Mass-spectrometric value.

‡ The McMaster standard speleothem 76001 has been dated many times by alpha-counting¹³. The mass-spectrometric dates were well within the error of the alpha-counting dates.

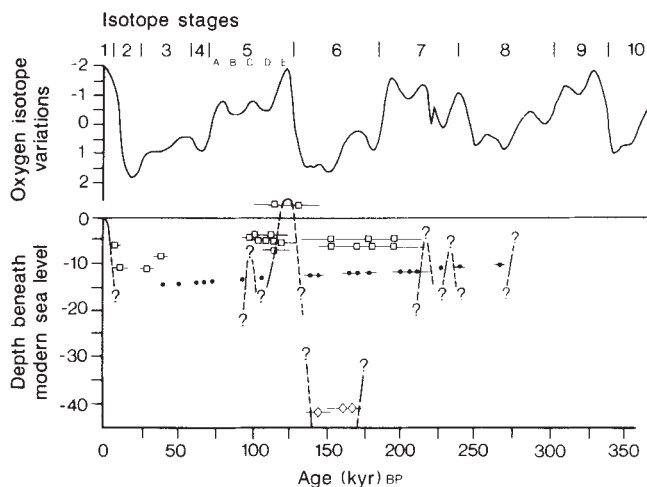


FIG. 3 Pleistocene sea-level curve for Bahamas from speleothem dates, correlated with the oxygen isotope record^{6,14}. The slope on which the speleothem dates are plotted is the maximum estimated tectonic subsidence rate² of 1 m per 50 kyr. The mass-spectrometric dates from this study are shown as solid circles, alpha-counted dates from Gascoyne *et al.*² as open diamonds and those from Harmon *et al.*^{1-3,7} as open squares.

Banks (estimated² at ~1 m per 50 kyr) and the difficulty of determining sea-level rise because of the erosion, the hiatuses indicate periods of sea level higher than -10 to -15 m at >280 kyr, 235-230 kyr, 220-212 kyr, 133-110 kyr, 100-97 kyr and <39 kyr (Fig. 3). The four rather short-lived peaks, at around 230, 215, 125 and 100 kyr BP, are generally in agreement with the high-sea-stand chronology inferred from oxygen isotope stratigraphy of oceanic foraminiferal cores⁶ and do not support the recent suggestion⁹ that the chronology of the marine foraminiferal record is questionable. However, the isotope variations in foraminifera are affected by other factors in addition to ice volume; sea-level changes follow the timing and general pattern of the foraminiferal isotopic variations but the magnitude of variations need not be directly comparable¹⁰. The Bahamas flowstone DWBAH sample shows no record of the expected (that is, to -15 m) sea-level rise in isotope stages 5a and 7a, although deeper speleothem may record rises to lower levels at these times. This sample also shows no evidence of the rise at ~40 kyr reported by Mylroie and Carew¹¹, but any part of it younger than 39 kyr has been lost to erosion. Further mass-spectrometric dating of submerged speleothem from various depths in the Bahamas and elsewhere should permit the construction of palaeo-sea-level curves of much greater precision than any that exist at present.

This study has shown that isotope-dilution mass-spectrometric U-series dating of speleothem is possible to at least 400 kyr BP on samples with U content as low as 0.08 p.p.m. Speleothem with high U content (>3 p.p.m.) should permit dating approaching 600 kyr BP. Where ²³⁴U/²³⁸U ratios can be shown to have remained constant, mass spectrometry will also give higher-precision age estimates beyond 600 kyr (ref. 12). As in alpha counting, the presence of detrital Th creates the problem associated with common ²³⁰Th. For the samples analysed here, however, with ²³⁰Th/²³²Th activity ratios greater than 100:1, this problem was negligible. □

8. Ivanovich, M. & Harmon, R. S. (eds) *Uranium Series Disequilibrium* (Oxford University Press, New York, 1982).
9. Winograd, I. J., Szabo, B. J., Coplan, T. B. & Riggs, A. C. *Science* **242**, 1275-1280 (1988).
10. Chappell, J. & Shackleton, N. J. *Nature* **324**, 137-140 (1986).
11. Mylroie, J. E. & Carew, J. L. *Quat. Sci. Rev.* **7**, 55-64 (1988).
12. Gascoyne, M., Ford, D. C. & Schwarcz, H. P. *Earth Surface Processes and Landforms* **8**(6), 557-568 (1983).
13. Harmon, R. S., Ku, T.-L., Matthews, R. K. & Smart, P. L. *Geology* **7**, 405-409 (1979).
14. Emiliani, C. J. *Geology* **63**, 538-578 (1955).

ACKNOWLEDGEMENTS. This research was supported by Natural Sciences and Engineering Research Council of Canada (A.P.D., D.C.F., R.McN. and H.P.S.).

Hatching asynchrony and reproductive success in the blackbird

Robert D. Magrath

Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK

A GENERAL problem of parental investment is that resources required to nourish offspring may be unpredictable at the time fertilization occurs¹⁻⁵. Lack¹ suggested that the best strategy for altricial birds faced with such unpredictability might be to lay an optimistic clutch size and to begin incubation before laying is complete. This causes the eggs to hatch asynchronously, which may allow the brood to be reduced to the optimum size by the quick elimination of the youngest nestling if food is scarce. If females hatched their eggs synchronously, food shortage could lead to poor growth of the whole brood if the nestlings are then equally competitive. Lack predicted that synchronous and asynchronous broods would consequently be equally productive in times of plenty, but asynchronous broods would be better when food is scarce^{1,3}. By manipulating both food availability and brood hierarchies, and following offspring survival after fledging, I show that in blackbirds (*Turdus merula*) asynchronous broods are more productive when food is scarce, but not when food is abundant.

Several hypotheses have been advanced to explain the normal hatching pattern of altricial birds, which is often asynchronous⁶, and it seems that different hypotheses are most plausible for different species⁶⁻¹¹. But there is still considerable debate as to the relevance of Lack's hypothesis to any species¹². Recent field experiments to test Lack's hypothesis that hatching asynchrony facilitates adaptive brood reduction are unsatisfactory because (1) food abundance and hatching pattern have not been simultaneously manipulated, and (2) juvenile survival has not been followed after fledging. These are serious limitations because Lack predicted that hatching asynchrony will be beneficial only when food is scarce, and fledging success can be a biased estimate of recruitment into the breeding population¹³⁻¹⁵. The latter problem is acute since asynchronous broods may fledge fewer, higher quality, young when food is scarce¹⁶. Here I present the first experimental test of Lack's hypothesis which addresses these problems.

I studied a population of the blackbird in the University Botanic Garden, Cambridge, England, from 1985 to 1988. The birds are monogamous, usually lay clutches of 3 or 4 eggs, and commonly raise 1 or 2 broods in the breeding season, which extends from March to July. Both parents feed the nestlings, which fledge when about 14 days old. The young are fed for at least 2 weeks after leaving the nest, but are independent by 4 weeks¹⁷. Hatching is usually asynchronous, with the last nestling hatching from the last-laid egg and suffering a reduced growth rate and increased chance of starvation (or disappearance) than its siblings. The incidence of such brood reduction varies among years and within breeding seasons, and is higher in larger broods, suggesting that food quite often can be limiting (manuscript in preparation). The nestlings are fed primarily on earthworms¹⁷,

Received 3 March; accepted 8 May 1989.

1. Harmon, R. S., Schwarcz, H. P. & Ford, D. C. *Quat. Res.* **9**, 205-218 (1978).
2. Gascoyne, M., Benjamin, G. J., Schwarcz, H. P. & Ford, D. C. *Science* **205**, 806-808 (1979).
3. Harmon, R. S. *et al. Nature* **289**, 481-483 (1981).
4. Schwarcz, H. P. in *Handbook of Environmental Geochemistry. The Terrestrial Environment, Part B* (eds Fritz, P. & Fontes, J.-C.) 271-303 (Elsevier, Amsterdam, 1986).
5. Edwards, R. L. *et al. Earth Planet. Sci. Lett.* **81**, 175-192 (1986/87).
6. Imbrie, J. *et al. in Milankovitch and Climate* (eds Berger, A. L. *et al.*) 269-305 (Reidel, Dordrecht, 1984).
7. Harmon, R. S. *et al. Palaeogeog. Palaeoclimat. Palaeoecol.* **44**, 41-70 (1983).

which show a general seasonal decline in availability, with reduced activity in dry periods¹⁸, during which nestling starvation is common¹⁷.

I mimicked variation in the hatching pattern of the birds in 1986 and 1987 by swapping recently-hatched nestlings among nests to produce broods of four which were either 'synchronous', where the nestlings were similar in size, or 'asynchronous', where the nestlings differed in size by an amount similar to the more asynchronous natural broods (Table 1). The nestlings were readily accepted by their foster parents. In 1986, I placed a nutritionally balanced¹⁹ food supplement in the territories of some pairs. It was placed on the ground twice daily, from just before hatching until 2 weeks after fledging. Most parents started using the food quickly, and delivered some of it to the nestlings; the others were excluded from all analyses. Within the constraints of hatchling availability, experimental treatments were alternated through each season. Nestlings were weighed and individually colour-ringed at day 8, and territories were searched for survivors 1 and 2 weeks after fledging. Records were also kept of all birds seen in regular searches of the garden, until the breeding season following hatching.

The mean growth rate of nestlings was higher, and there was no seasonal decline, in broods provided with a food supplement in 1986, suggesting that the seasonal decline in growth rate in unsupplemented broods was indeed due to a worsening natural food supply (Table 2). The mean growth of the (unsupplemented) broods in 1987 was also low, and the incidence of brood reduction was substantial (Table 2). I therefore defined 'good' conditions as occurring if a food supplement was provided, or it was early in 1986; other broods suffered 'poor' conditions.

The crucial test of Lack's hypothesis is to compare the productivity of experimental broods under good and poor conditions (Fig. 1). The mean productivity of both brood types was lower in poor conditions, but the productivity of successful asynchronous broods did not differ statistically between good and poor conditions, whereas successful synchronous broods produced significantly fewer young in poor conditions. The productivity of asynchronous broods was greater than synchronous broods in poor conditions, as Lack predicted. In good conditions the productivity of the two brood types did not differ statistically, again as Lack predicted, but the sample size in good conditions was smaller than in poor conditions, and therefore the chance of detecting a real difference was lower. The figure suggests that synchronous broods might be better in good conditions. If there is a cost to asynchrony in good conditions then the optimal asynchrony may depend on the relative frequency of good and poor conditions. Possible costs of asynchrony include premature brood reduction²⁰⁻²², a reduced hatching success²⁰ and less synchronous fledging³. In blackbirds an optimum may well lie between the two experimental groups, as does the mean natural asynchrony. Broods failing completely were excluded from the analysis because these losses were caused mostly by predation (26/33 failures), which is not directly relevant to assessing Lack's

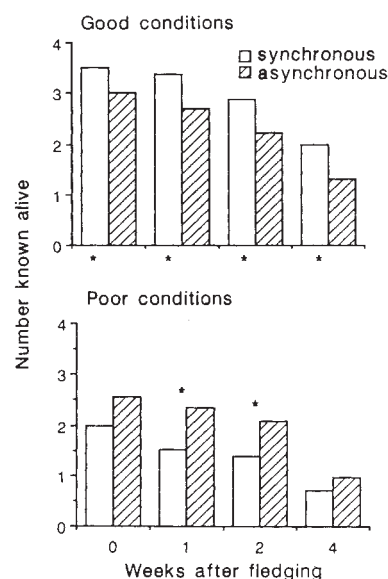


FIG. 1 Mean number of nestlings known to have fledged, from synchronous (plain bars) and asynchronous (hatched bars), and to have survived to 1, 2 and 4 weeks from fledging, in broods from which at least one nestling was known to have fledged. An asterisk below a bar indicates a statistically significant difference between good and poor conditions; one above columns shows a significant difference between synchronous and asynchronous broods (Mann-Whitney *U* Test; $P < 0.05$). Lack of an asterisk indicates that the difference was not significant ($P > 0.05$). The most accurate measures of brood productivity are at 1 and 2 weeks, when territories were searched for dependent young; the estimate of survival to 4 weeks could be influenced by dispersal. Loglinear analyses²⁹ of death records and sightings of 748 nestlings from 1985 to 1987 showed that weight at 8 days influenced survival up until 4 weeks after fledging, but that breeding birds were a random sample of those alive at 4 weeks. Most mortality of birds less than a year old had occurred by about 4 weeks after fledging; most differential mortality by 2 weeks after fledging (manuscript in preparation). Sample sizes: good conditions, $n=8$ synchronous broods, 13 asynchronous broods; poor conditions, $n=21$ synchronous broods, 25 asynchronous broods.

TABLE 2 Defining 'good' and 'poor' conditions

Environmental conditions	<i>n</i>	Mean weight at 8 days (g) ($\pm$ s.e.)	Starve or disappear (percentage of broods involving ≥ 1 nestling)	Starve, disappear or poor growth (percentage of broods involving ≥ 1 nestling)
Good				
Food; early; 1986	8	59.9 (± 1.49)	12.5	12.5
Food; late; 1986	6	59.0 (± 2.19)	0.0	0.0
No food; early; 1986	13	59.2 (± 0.38)	7.7	7.7
Poor				
No food; late; 1986	15	50.9 (± 1.91)	46.7	66.7
No food; 1987	37	53.8 (± 0.95)	35.1	43.2

The food supplement consisted of wholemeal bread, rolled oats, grated cheese, soaked currants and mealworms. Weight was standardized to 8.0 days by using the logistic growth curve fitted^{27,28} to repeated measures of weight of nestlings that subsequently fledged; thus broods did not have to be disturbed repeatedly to measure each nestling at exactly 8.0 days. Nestlings which 'disappeared' did so singly before day 8. Nestlings which weighed less than 45 g at day 8 were considered to have grown 'poorly', because there was a rapid decline in the probability of fledging below that weight. 'Early' broods were essentially first broods, 'late' broods were generally second or repeat broods; they represented the two peaks in laying dates. In 1987, the experiment was carried out primarily on late broods because broods hatching in the first 3 weeks of the breeding season were involved in another experiment. One-way analysis of variance of the mean weights of nestling in broods at day 8, with a Student-Newman-Keuls range test, showed that there was no difference amongst the three groups in 'good' conditions, whereas the two groups in 'poor' conditions were both lighter, but not different from each other (overall, $F_{4,74} = 6.64$; $P = 0.0001$). *n*, Number of broods.

TABLE 1 Hatching hierarchies of experimental and natural broods

	Experimental (mean $\pm$ s.d.)		Natural (mean $\pm$ s.d.; range)
	Synchronous (<i>n</i> = 42)	Asynchronous (<i>n</i> = 58)	(clutch of 4; all hatch) (<i>n</i> = 64)
Age difference (days)	0.49 (± 0.25)	1.25 (± 0.39)	0.87 (± 0.38 ; 0.0-1.8)
Weight difference (g)	1.18 (± 1.06)	5.17 (± 1.78)	3.38 (± 1.95 ; 0.6-8.7)
Tibia difference (mm)	0.99 (± 0.62)	3.30 (± 0.98)	2.26 (± 1.12 ; 0.8-5.4)

Nestlings were individually marked, at or soon after hatching, by clipping specific tufts of down. All experimental broods included at least one adopted nestling—synchronous 2.0 (1-4); asynchronous 1.6 (1-4) mean (range). Most natural broods shown were manipulated after hatching to become experimental broods. *n*, Number of broods.

hypothesis concerning brood reduction and sibling competition. (The degree of asynchrony could influence the probability of predation if nestlings beg louder in one brood type; if so the effect would have to be balanced against the direct effect of competition.) There was also no difference in the proportion of synchronous and asynchronous broods suffering total failure, nor in the proportion of failures caused by desertion (13/42 synchronous and 20/58 asynchronous broods failed, $G_1 = 0.14$, $P = 0.71$; of these 3/13 and 4/20 broods respectively were deserted).

The reason that asynchronous broods should do better than synchronous broods in poor conditions, Lack argued, is that brood reduction will be more efficient in asynchronous broods³. In support of this idea, I found that if brood reduction occurred, it did so earlier in asynchronous broods: 9/14 compared with 2/14 nestlings starved or disappeared before day 8 in asynchronous compared with synchronous broods ($G_1 = 7.79$; $P < 0.01$). Because nestlings died later in synchronous broods, energy was wasted on nestlings which subsequently died, as has been found in the jackdaw, *Corvus monedula*²³. The earlier brood reduction in asynchronous broods was of benefit to the survivors. The mean weight of surviving nestlings in broods in which reduction occurred was heavier in asynchronous (55.36 ± 1.54 , $n = 13$ broods) than synchronous broods (49.57 ± 2.03 , $n = 9$ broods; mean $\pm$ s.e.; $t = 2.31$, d.f. = 20, 1-tailed $P = 0.02$), but in broods in which no reduction occurred nestlings grew at similar rates in asynchronous (57.15 ± 0.86 , $n = 34$ broods) and synchronous broods (56.33 ± 1.01 , $n = 23$ broods; mean $\pm$ s.e.; $t = 0.61$, d.f. = 55, 2-tailed $P = 0.54$).

Not all observations were consistent with Lack's original formulation of the mechanism of brood reduction. The range of growth rates amongst synchronous broods approached that of asynchronous broods (Table 3), contrary to Lack's prediction that nestlings which were of similar size earlier in the nestling period would be equally competitive, and would therefore grow at similar rates, whatever the feeding conditions. It has also been reported from other species that differences in growth can occur between similarly-aged nestlings²³⁻²⁶, so competitive hierarchies amongst nestlings in synchronous broods may differ quantitatively, and not qualitatively, from those in asynchronous broods. It seems likely that in natural broods there will always be some differences in behaviour or size which will influence competition^{7,12,23}, as there were in experimental broods. Direct observations of the role of parents and nestlings in the allocation of food in synchronous and asynchronous broods should help clarify the mechanisms of brood reduction in small passerines.

Although the relevance of Lack's hypothesis to other species must be demonstrated, it is premature to conclude, as do Amundsen and Stokland¹², that "the brood reduction hypothesis does not provide an explanation of hatching asynchrony in general". Such a conclusion can only follow from other experimental tests in which food availability is quantified and juvenile survival

determined. A challenging theoretical and empirical problem is to show the optimum hatching asynchrony, given known environmental conditions, and thus to make Lack's hypothesis quantitative rather than qualitative. □

Received 6 February; accepted 11 May 1989.

1. Lack, D. *Ibis* **89**, 302-352 (1947).
2. Martin, T. E. *A. Rev. Ecol. Syst.* **18**, 453-487 (1987).
3. Lack, D. *Ecological Adaptations for Breeding in Birds* (Methuen, London, 1968).
4. Stephenson, A. G. *A. Rev. Ecol. Syst.* **12**, 253-279 (1981).
5. Perrigo, G. *Anim. Behav.* **35**, 1298-1316 (1987).
6. Clark, A. B. & Wilson, D. S. *Q. Rev. Biol.* **56**, 253-277 (1981).
7. Mock, D. W. in *Infanticide: Comparative and Evolutionary Perspectives* (eds Hausfater, G. & Hrdy, S. B.) 3-30 (Aldine, New York, 1984).
8. Bryant, D. M. *Ibis* **120**, 16-26 (1978).
9. Hahn, D. C. *Anim. Behav.* **29**, 421-427 (1981).
10. Stinson, C. H. *Evolution* **33**, 1219-1225 (1979).
11. Cash, K. J. & Evans, R. M. *Behav. Ecol. Sociobiol.* **18**, 413-418 (1986).
12. Amundsen, T. & Stokland, J. N. *J. anim. Ecol.* **57**, 329-344 (1988).
13. Perrins, C. M. *J. anim. Ecol.* **34**, 601-647 (1965).
14. Gustafsson, L. & Sutherland, W. J. *Nature* **335**, 813-815 (1988).
15. Pettifor, R. A., Perrins, C. M. & McCleery, R. H. *Nature* **336**, 160-162 (1988).
16. Slagsvold, T. *J. anim. Ecol.* **55**, 1115-1134 (1986).
17. Snow, D. W. *Ibis* **100**, 1-30 (1958).
18. Edwards, C. A. & Lofty, J. R. *Biology of Earthworms* (Chapman & Hall, London, 1972).
19. Meaden, F. *A Manual of European Bird Keeping* (Blandford, Poole, 1979).
20. Slagsvold, T. *Ornis. Scand.* **16**, 81-87 (1985).
21. Skagen, S. K. *Ecology* **68**, 1747-1759 (1987).
22. Skagen, S. K. *Auk* **105**, 78-88 (1988).
23. Gibbons, D. W. *J. anim. Ecol.* **56**, 403-414 (1987).
24. Bengtsson, H. & Rydén, O. *Behav. Ecol. Sociobiol.* **12**, 243-251 (1983).
25. Greig-Smith, P. *J. Zool.* **205**, 453-465 (1985).
26. Moreno, J. *Ornis. Scand.* **18**, 302-309 (1987).
27. Ricklefs, R. E. *Avian Biology Vol. VII* (eds Farner, D. S., King, J. R. & Parkes, K. C.) 1-82 (Academic, New York, 1983).
28. Ross, G. J. S. *Maximum Likelihood Program* (Rothamsted Experimental Station, 1980).
29. Feinberg, S. E. *The Analysis of Cross-classified Categorical Data* (MIT, Cambridge, 1977).

ACKNOWLEDGEMENTS. I thank Nick Davies for discussion, advice and encouragement in all aspects of the study, and Tim Clutton-Brock, André Desrochers, Andy Horn, Marty Leonard for valuable comments on the manuscript. For permission to work in the Botanic Garden, I thank C. D. Piggett and P. Orriss. The experiment was done under licence from the Nature Conservancy Council.

Transgenic mice with inducible dwarfism

Emiliana Borrelli^{*‡}, Richard A. Heyman^{*}, Carlos Arias[†], Paul E. Sawchenko[†] & Ronald M. Evans^{*}

^{*} Gene Expression Laboratory, Howard Hughes Medical Institute and

[†] Laboratory of Neuronal Structure and Function, The Salk Institute for Biological Studies, La Jolla, California 92037, USA

THE pituitary gland, composed of the anterior, intermediate and posterior lobe, represents a principal regulatory interface through which the central nervous system controls body physiology¹. The ontogeny of the growth hormone (GH) and prolactin (Prl) producing cells of the anterior pituitary has been analysed in transgenic mice, using the thymidine kinase obliteration system (TKO)^{2,3}. Cells expressing the herpes virus 1 thymidine kinase (HSV1-TK) gene acquire pharmacological sensitivity to synthetic nucleosides such as FIAU (1-(2-deoxy-2-fluoro- β -D-arabinofuranosyl)-5-iodouracil), whose metabolites kill dividing cells. Consequently we created transgenic mice carrying the HSV1-TK gene under the control of either the rat growth hormone or the rat prolactin promoter. If transgenic mice expressing HSV1-TK in somatotropes (GH-producing cells) are treated with FIAU, they develop as dwarfs. The anterior pituitary in these animals is nearly devoid of both somatotropes and lactotropes (Prl-producing cells). By contrast, transgenic mice expressing HSV1-TK in the lactotropes, treated with FIAU, have anatomically and histologically normal pituitaries. Because toxicity depends on cell division, we conclude that Prl expression and lactotrope differentiation are post-mitotic

TABLE 3 Differences in growth in experimental broods in which all four nestlings were alive at day 8

Environmental conditions	Synchronous	Asynchronous
Good	5.18 (± 2.6 ; $n = 10$)	8.79 (± 3.7 ; $n = 15$)
Poor	12.53 (± 7.2 ; $n = 20$)	14.07 (± 8.9 ; $n = 22$)

Difference between fastest and slowest growing nestlings at day 8. Weights were standardized to day 8 (see Table 2 legend), and the mean difference ($\pm$ s.d.) between the heaviest and lightest is shown (n , number of broods). Two-way analysis of variance of log-transformed data showed that there was a difference between synchronous and asynchronous broods ($F_{1,63} = 4.35$, $P = 0.04$) but there was indication of an interaction ($F_{1,63} = 3.71$, $P = 0.06$) suggesting that the difference lay primarily in good conditions, when the mean range of growth was lower ($F_{1,63} = 20.06$, $P < 0.001$).

[‡] Present address: Unite 184 de Biologie Moléculaire et de Génie Génétique de l'INSERM, 11 Rue Humann, 67085 Strasbourg Cedex, France

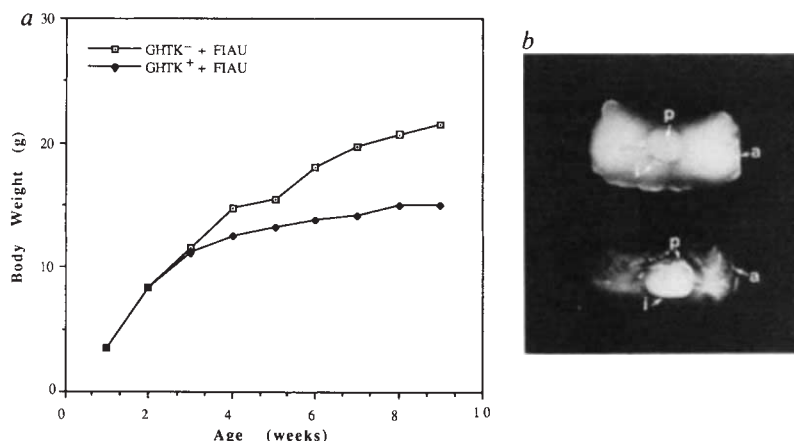


FIG. 1 Effect of FIAU treatment on animals expressing HSV1-TK in the somatotropes. **a**, Body weight of the GHTK transgenic versus non-transgenic littermate, both treated with FIAU. Each point is the mean of the weights of at least 10 animals. **b**, Comparison of the pituitary glands of a normal non-transgenic mouse: top, with a transgenic GHTK; bottom, both treated for 8 weeks. p, posterior lobe; i, intermediate lobe; a, anterior lobe. **METHODS.** Normal females were mated with transgenic males or transgenic females were mated with normal males. The pregnant females were implanted subcutaneously at day 5 of gestation with an Alzet pump delivering FIAU at a rate of 1 mg per day. The newborns were injected daily from the second day of life, with 50 μ l of FIAU in PBS corresponding to 0.04 mg $\cdot$ g⁻¹ per day.

events. These results indicate that both somatotropes and lactotropes derive from a common GH-expressing stem-somatotrope. Unexpectedly, the stemsomatotrope is still present in the adult animal and is capable of repopulating the pituitaries of treated animals with mature GH and Prl producing cells.

In rodents the pituitary gland is composed of the anterior, intermediate and posterior lobe, the former of which comprises five cell types: the somatotropes, lactotropes, thyrotropes, gonadotropes and corticotropes, distinguished by synthesis, release, and storage of a specific hormone.

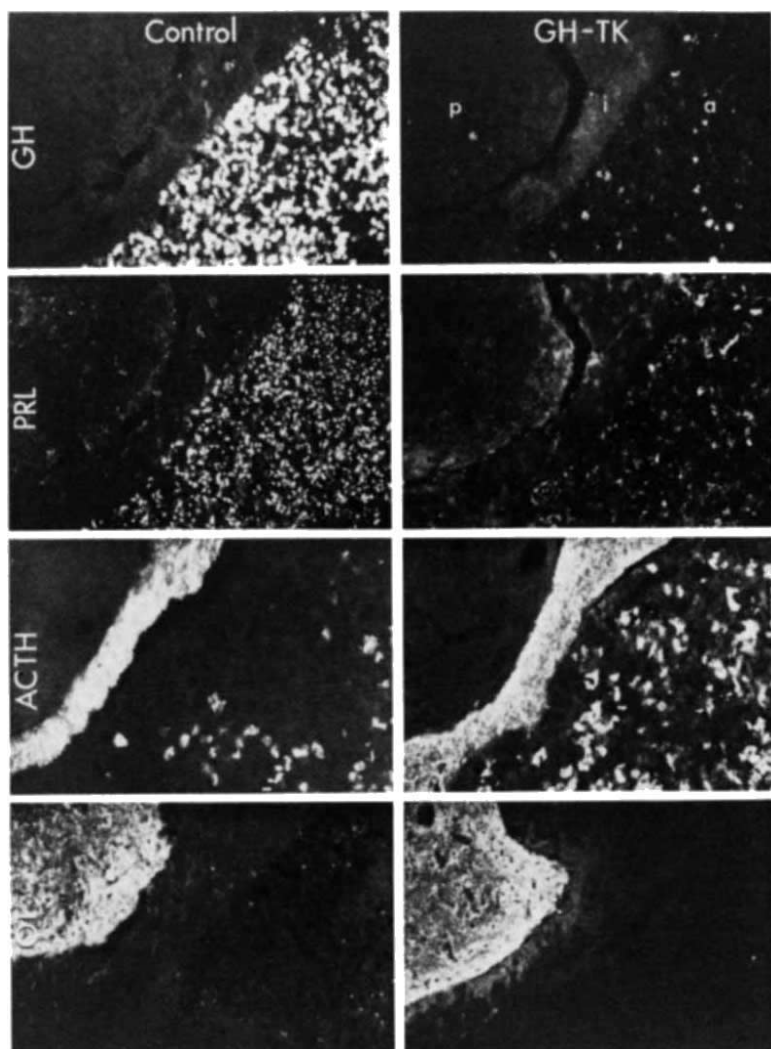
To establish transgenic pedigrees for analyses of the

somato/lactotrope lineage, a 1.5-kilobase (kb) fragment (*HincII*-*NcoI*) of the HSV1-TK gene was fused to a 320-base pair (bp) fragment (*KpnI*-*XhoI*/-310, +8) and to a 3,033-bp fragment (*BamHI*-*HindIII*/-3,000, +33) of the rat GH or Prl promoters, respectively. The resulting GHTK and PrlTK constructs were microinjected into the male pronucleus of fertilized eggs⁶.

Previous studies have established that these promoters contain enhancer elements that allow for their tissue specific expression in the appropriate cell types in the anterior pituitary⁷⁻¹⁰. Various transgenic pedigrees were obtained each of which expressed HSV1-TK selectively in the anterior pituitary. Enzymatic assay

FIG. 2 Effect of FIAU treatment on pituitary cell and fiber types in GHTK transgenic mice. Comparison of indirect immunofluorescence-staining patterns in transgenic (GHTK) mice treated with FIAU from postnatal day 2 to 8-10 weeks of age with non-transgenic (Control) littermates. A marked decrease in the number of cells expressing GH- and Prl-immunoreactivities in the anterior lobe (a) is evident. ACTH-stained cells in the anterior lobe appear more numerous in the anterior lobe of treated animals, presumably as a consequence of cell loss in the anterior lobe as a whole. Staining for ACTH in the intermediate lobe (i) is similar in the two groups, as are oxytocinergic (OT) terminals in the posterior lobe (p). All micrographs $\times 90$.

METHODS. Mice were anaesthetized with chloral hydrate and perfused transcardially with 10% neutral buffered formalin (Richard Allan). Pituitaries were removed and post-fixed overnight at 4 $^{\circ}$ C in the same fixative used to perfuse the animals with 10% sucrose added. Frozen sections (20- μ m thick) were cut on a sliding microtome. Pituitary sections were processed for indirect immunofluorescence. The antisera used were monkey antimouse growth hormone (1:80,000 dilution, from Y. Sinha, Whittier Institute, California) rabbit antihuman prolactin (1:1,000, from Middlesex Turnpike, Massachusetts) rabbit antihuman corticotropin (1:1,000 from Wylie Vale, Salk Institute, California) rabbit antirat oxytocin (1:2,000 from S. Vandesande and K. Dierckx, Gent, Belgium). The secondary antisera used were raised in goat and were fluorescein- or rhodamine-labelled against monkey IgG (1:100 from Cappel, Malvern, Pennsylvania), rabbit IgG (1:200 from Tago, Burlingame, California), rabbit antirat TSH (1:2,000), LH (1:5,000) and FSH (1:2,000) from the National Pituitary and Hormone Distribution Program of the NIADDK. Stained sections were mounted on glass slides, coverslipped with 50% glycerol in 0.4 M potassium bicarbonate buffer (pH 8.6). The fluorescence microscope used was a Leitz Dialux 20.



for HSV1-TK activity¹¹ revealed in each case the tissue specific expression of this gene product (Table 1). To initiate pituitary ablation, the 388 (GHTK) and 666 (PrITK) line were chosen as they were among the ones with the higher HSV1-TK activity (see Table 1). GHTK and PrITK transgenic and nontransgenic embryos were treated with FIAU (see Fig. 1 legend) from embryonic day 5 through birth, and the neonates from 48 h after birth to 8–10 weeks of age. All transgenic animals gained weight at rates comparable to those of their nontransgenic littermates during the first 2 weeks of age. Over the next 2 weeks a slight decrease in growth was noticed in the transgenic mice. At 4 weeks of age, the weight gain in GHTK transgenics treated with FIAU ceased, whereas that of the Prl animals and the nontransgenic littermates treated with FIAU progressed normally (Fig. 1).

Macroscopically, pituitaries from transgenic GHTK animals were atrophied and weighed only 200 µg compared to 1 mg from the nontransgenic littermates. Although the posterior and intermediate lobes appeared normal, the anterior lobe was poorly developed and nearly translucent (Fig. 1). By contrast, pituitaries from PrITK animals treated with FIAU for 10 weeks appeared normal. Immunohistochemical localization^{10,12} of markers for each of the principal cell and fibre types in the pituitary was carried out. This involved concurrent dual localization of GH- and Prl-immunoreactivity, as well as staining of neighbouring sections for other pituitary hormones including ACTH, FSH, LH, TSH, oxytocin and vasopressin. The most obvious alterations seen in FIAU-treated GHTK animals was a near-total depletion of cells displaying GH and/or Prl immunoreactivity (Fig. 2). By contrast, cells expressing ACTH, TSH, FSH, and LH in the anterior lobe appeared unaffected in the treated transgenics and controls; the density of these cell types was greater in the treated GHTK mice and may be due to the absence of somatotropes and lactotropes thereby bringing the other cell types physically closer (see Fig. 2 ACTH). Oxytocin- and vasopressin-immunoreactive fibers in the posterior lobe appeared unaffected, as did ACTH-stained cells in the intermediate lobe (Fig. 2).

Unexpectedly, a similar immunohistochemical analysis of treated and control animals from two PrITK pedigrees failed to

reveal any effect on somatotrope or lactotrope populations. Their pituitaries were anatomically and physiologically normal after 10 weeks of treatment. Because toxicity is dependent on cell division, we conclude that prolactin expression and lactotrope differentiation are post-mitotic events. The small pool of surviving cells (<1% of normal) is probably a consequence of incomplete toxicity, although a minor pathway of differentiation cannot be excluded. These results of the GHTK and PrITK mice suggest that a replicating GH-expressing cell (a stem-somatotrope) serves as a common precursor to the final pool of nondividing, mature somatotropes and lactotropes.

The TKO technique can be used to explore the frequency of which post-mitotic GH and Prl expressing cells are deposited into the pituitary by delaying onset of drug treatment. When treatment was initiated at 2, 5 or 10 days after birth, a graded increase in somatotrope numbers in the pituitaries of treated animals was observed when compared with the ones where treatment started at day 2 (Fig. 3), although at each of these times the lactotrope population remained completely repressed. Furthermore, the group of animals treated from day 10 were normal in size whereas the animals treated from day 2 or day 5 were dwarfs. These results suggest that terminal somatotrope differentiation begins early and is a progressive event. These results also suggest that post-mitotic somatotropes cannot progress toward lactotropes.

To investigate the regenerative potential of pituitary stem cells, animals initially treated for 8 weeks were allowed to recover from FIAU treatment for 3 or 6 weeks before sacrifice. The immunohistochemical analysis demonstrates a clear rebound of somatotropes in these two time periods (Fig. 4) as well as a 1.5-g and 4-g increase in weight, respectively. Partial rebound

TABLE 1 HSV1-Thymidine kinase activity in the tissue of the GHTK line and the PrITK line

		HSV1-TK activity (pmol per min per mg protein)			
GHTK line		PrITK lines			
388	Pit	13	663	Pit	14
	Liv	<0.03		Liv	<0.03
	Kid	<0.03		Kid	<0.03
	Lu	<0.03		Lu	<0.03
	Brain	<0.03		Brain	<0.03
	Cereb	<0.03		Cereb	<0.03
	Heart	<0.03		Heart	<0.03
385	Pit	5	666	Pit	100
	Liv	<0.03		Liv	<0.03
	Kid	<0.03		Kid	<0.03
	Lu	<0.03		Lu	<0.03
	Brain	<0.03		Brain	<0.03
	Cereb	<0.03		Cereb	<0.03
	Heart	<0.03		Heart	<0.03

Homogenates of tissues from GHTK and PrITK animals were prepared and assayed for HSV1-TK activity. Mean values of three separate experiments are shown. Tissues from transgenic GHTK and PrITK animals were homogenized and centrifuged at 15,000g for 15 min at 4 °C. The supernatant was assayed for HSV1-TK activity using ³(H)-acyclovir as a substrate, monitoring its conversion to ³(H)-acyclovir-mono, diphosphates and triphosphates as described previously¹¹. The phosphorylated products were separated from substrate on TLC (PEI-cellulose). Quantitation of the results was obtained by determination of the radioactivity.

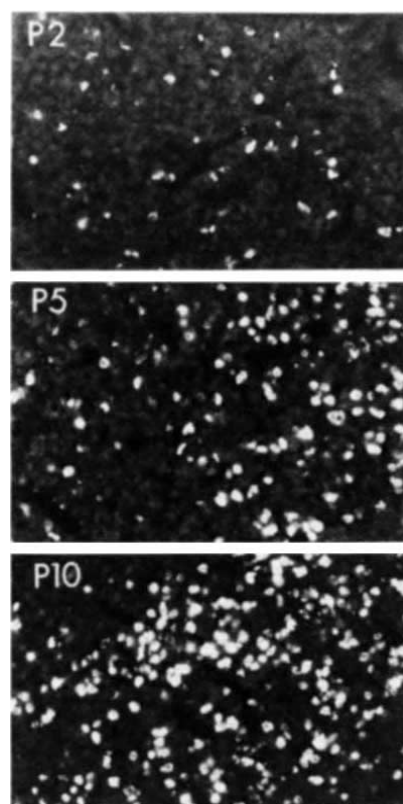


FIG. 3 Differential effect on anterior pituitary GH-immunoreactivity as a function of the age at which substrate treatment was initiated in GHTK transgenic mice. Mice were treated continuously (until killed) with FIAU beginning on postnatal day 2, 5 or 10 (P2, P5, P10). A clearly graded effect on the number of GH-immunoreactive cells was apparent. Note that even in the P10 group, the number of somatotropes remains considerably lower than that seen in nontransgenic controls (see Fig. 2). All micrographs $\times 75$.

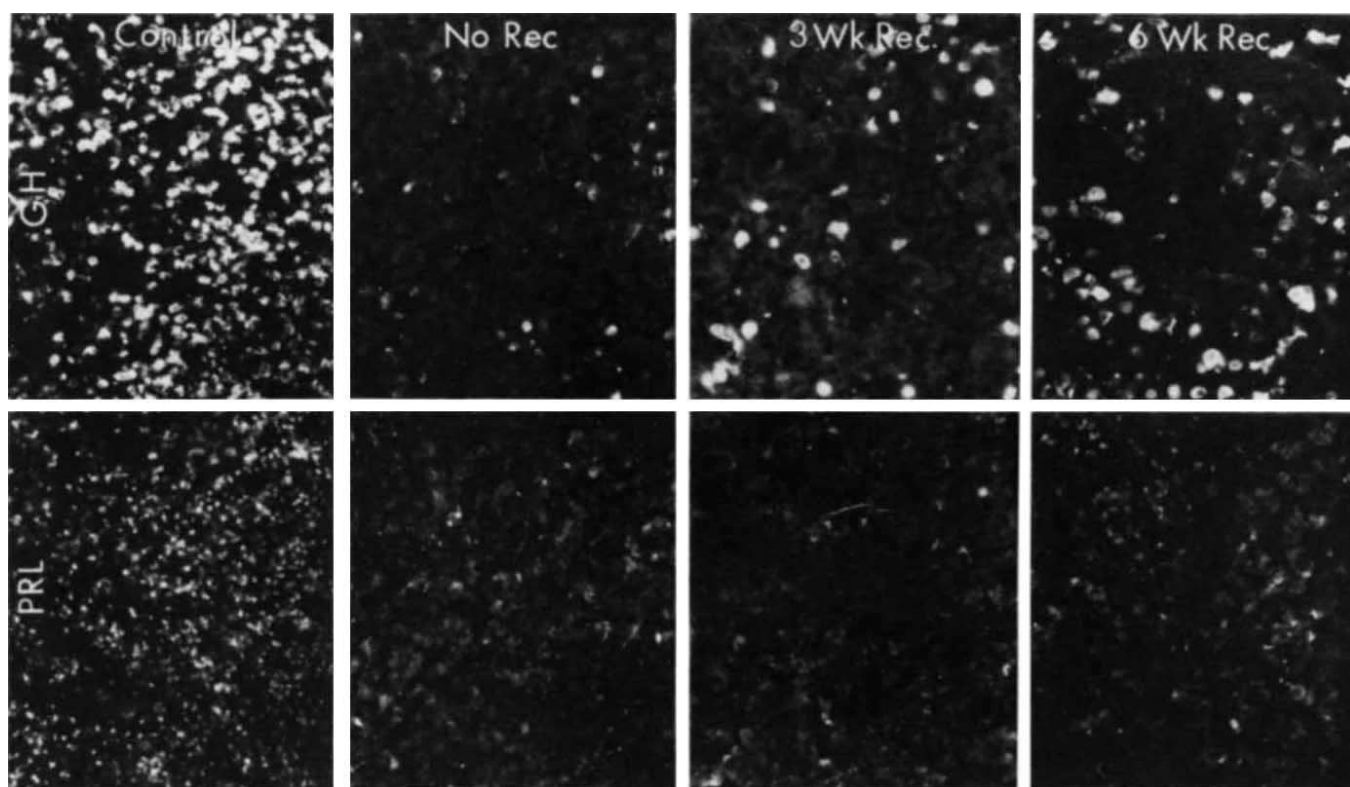


FIG. 4 The effect of cessation of substrate treatment on GH- and PRL-immunoreactivity in GH-TK transgenic mice. Mice were treated from postnatal day 2 to 8 weeks of age, at which time treatment was discontinued for 3 (3 Wk Rec.) or 6 (6 Wk Rec.) weeks. Comparison with a FIAU-treated, nontransgenic littermate (Control) and a continuously treated transgenic

littermate (No Rec.) are provided. Individual sections through the anterior lobe stained for GH (top row) and PRL (bottom) are shown for each group. GH cells increase in number and size after withdrawal of substrate treatment. Any 'recovery' of PRL cells is not apparent until 6 weeks following termination of FIAU treatment. All micrographs $\times 150$.

of lactotropes was also observed but only after 6 weeks of recovery (Fig. 4). Together, the results of this study suggest that a dividing and potentially self-renewing stem-somatotrope serves as a common precursor to the terminally differentiated somatotrope and lactotrope populations. In a similar study¹³ mice transgenic for the diphtheria toxin gene fused to the GH promoter developed as dwarfs and completely lacked somatotropes. Although lactotropes were greatly decreased, the retention of a residual population led to a proposed dual pathway of lactotrope generation. The combination of our cytotoxic and rebound data, however, argues against this proposed binary pathway and instead supports a model of direct descent. □

Clonal anergy induced in mature $V\beta 6^+$ T lymphocytes on immunizing *Mls-1^b* mice with *Mls-1^a* expressing cells

Hans-Georg Rammensee, Ruth Kroschewski & Bernard Frangoulis*

Max-Planck-Institut für Biologie, Abteilung Immunogenetik, Corrensstrasse 42, D-7400 Tübingen, FRG

TOLERANCE to self-antigens has been shown to develop during ontogeny as a result of the clonal deletion of self-reactive T cells¹⁻⁴. Tolerance, or better 'nonresponsiveness', to specific antigens can also be induced in adult animals⁵⁻⁹ but the mechanism(s) involved are not well understood¹⁰. Most murine T-helper cells that express the $V\beta 6$ T-cell receptor gene segment are specific for *Mls-1^a* antigens³. We have therefore been able to use an anti- $V\beta 6$ monoclonal antibody to follow the fate of *Mls-1^a* specific T cells in adult *Mls-1^b* mice made specifically unresponsive to *Mls-1^a*. We show that the induced unresponsiveness is not due to clonal deletion, but rather to clonal anergy. The anergic $V\beta 6$ T-helper cells express IL-2 receptors and undergo limited blastogenesis *in vitro* upon stimulation, but do not produce IL-2, in marked contrast to $V\beta 6$ cells from naive mice. Our data appear to represent an *in vivo* correlate for the induction of anergy that has been observed in T-cell lines *in vitro*¹¹⁻¹⁵.

* Present address: Institut de Recherches sur les Maladies du Sang, Hôpital Saint-Louis, 2, Place Fournier, Paris, France

Received 7 March; accepted 24 April 1989.

1. Felig, P., Baxter, J. D., Brodaus, A. E. & Frohman, L. A. *Endocrinology and Metabolism* (McGraw-Hill, New York, 1987).
2. Borrelli, E., Heyman, R., Hsi, M. & Evans, R. M. *Proc. natn. Acad. Sci. U.S.A.* **85**, 7572-7576 (1988).
3. Heyman, R. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 2698-2702 (1989).
4. McLaren, C., Chen, M. S., Barhaiya, R. H., Buroker, R. A. & Oleson, F. B. in *Herpes Viruses and Virus Chemotherapy* (ed. R. Kono) 57-61 (Elsevier, Amsterdam, 1985).
5. Mansuri, M. M., et al. *J. med. Chem.* **30**, 867-871 (1987).
6. Hogan, B., Costantini, F. & Lacy, E. *Manipulating the Mouse Embryo: a Laboratory Manual* (Cold Spring Harbor, New York, 1986).
7. Nelson, C. et al. *Nature* **322**, 557-562 (1986).
8. Nelson, C., Albert, V. R., Elsholtz, H. P., Lu, L. I.-W. & Rosenfeld, M. G. *Science* **239**, 1400-1405 (1988).
9. Lira, S. A., Crenshaw, E. B. III, Glass, C. K., Swanson, L. W. & Rosenfeld, M. W. *Proc. natn. Acad. Sci. U.S.A.* **85**, 4755-4759 (1988).
10. Crenshaw, E. B. III et al. *Genes Dev.*, in the press.
11. Fyfe, J. A., Keller, P. M., Furman, P. A., Miller, R. L. & Elion, G. B. *J. biol. Chem.* **253**, 8721-8727 (1978).
12. Sawchenko, P. E. et al. *J. comp. Neurol.* **241**, 138-153 (1985).
13. Behringer, R. R., Mathews, L. S., Palmiter, R. D. & Brinster, R. L. *Genes Dev.* **2**, 453-461 (1988).

ACKNOWLEDGEMENTS. We thank Kevin Murakami, Deborah Anderson and Jenny Price for technical help, Drs Wylie Vale and Douglas Richman for discussions, Tony Oro and Stan Hollenberg for comments on the manuscript and Geof Rosenfeld for communicating unpublished information. We also thank Dr Larry Swanson and Donna Simmons for advice in preparing histologic material and to Elaine Stevens for typing the manuscript. R.M.E. is an investigator and R.A.H. is a Research Associate of the Howard Hughes Medical Institute. This work was supported by the Mathers Foundation, the Howard Hughes Medical Institute and the National Institutes of Health.

TABLE 1 Inhibition of the Mls-specific response and analysis of V β 6⁺CD4⁺ T cells in C3H mice pretreated with Mls-1^a antigen

Injection of Mls-1 ^a - expressing cells (days before MLR)	Mixed lymphocyte reaction [³ H]thymidine incorporation (c.p.m.) on stimulation with cells from the mouse strains:				Max. ratio† of response to Mls versus H-2	Percentage of lymph node cells which are:			
	C3H(<i>k; b</i>)*	C3H.CE(<i>k; a</i>)	AKR(<i>k; a</i>)	BALB/c (<i>d; b</i>)		Vβ6 ⁺	Vβ6 ⁺ among		
							Vβ6 ⁺ CD4 ⁺	CD4 ⁺	Vβ6 ⁺ CD4 ⁻
Expt 1									
no injection‡	11,971	380,741	244,607	96,033	4.4	11.6	6.9	13.4	4.7
day-9	8,208	358,165	227,505	68,219	5.8	10.9	6.4	13.0	4.5
	10,360	26,789	18,837	71,180	0.26	12.7	8.7	20.6	4.0
	12,231	44,531	37,103	64,094	0.62	14.7	10.2	21.0	4.5
Expt 2									
no injection	3,260	70,119	ND	19,065	4.2	10.2	6.9	14.5	3.2
day-12	2,301	9,114	ND	24,587	0.30	6.6	4.1	5.9	2.5
	1,915	3,410	ND	15,221	0.11	9.4	4.7	8.7	4.7

C3H/HeJ mice (*H-2^k*; *Mls-1^b*) were either left untreated or given one tail vein injection of 10⁷ C3H.CE-Ly-1^b (abbreviated C3H.CE; *H-2^d*, *Mls-1^a* [B. F., unpublished data]) spleen cells, suspended in 0.3 ml phosphate buffered saline. At the indicated time intervals after injection, lymph node cells were stimulated in a mixed lymphocyte reaction (MLR) using cells from the different strains shown. At the same time, an aliquot of each lymph node cell population was tested by two colour fluorescence analysis for expression of V β 6 and CD4 molecules (see legend to Fig. 1). For the MLR assay cells of inguinal, axillary, mesenteric and paraaortic lymph nodes of recipients or untreated mice were cultured in 96-well flat bottom microtitre plates (Greiner) in triplicates at 5 × 10⁵ cells per well (Expt 1) or 2 × 10⁵ cells per well (Expt 2) together with syngeneic or allogeneic stimulator cells (irradiated with 10 Gy from a ¹³⁷Cs source) at 10⁶ cells per well, in Alpha-MEM medium (Gibco) supplemented with 10% fetal calf serum, 2-mercaptoethanol, L-glutamine and antibiotics. After 3 days of culture, 37 kBq of [³H]thymidine (Amersham) were added. Cells were collected at day 4 of culture onto glass fibre filter mats. Radioactivity was determined in a liquid scintillation counter. Data are indicated as arithmetic mean of triplicates. Standard deviation was normally less than 15% of the mean. ND, not determined.

* *H-2* haplotype; *Mls-1* allele²⁶.

† Δ c.p.m. Mls; Δ c.p.m. H-2.

‡ Each line corresponds to one individual C3H mouse.

Lymphocytes from C3H/HeJ (*H-2^k*, *Mls-1^b*) mice produce a strong proliferative response against *H-2* matched, *Mls*-disparate cells from AKR/J (*Mls-1^a*) or C3H.CE (*Mls-1^a*) mice^{16,17}, (see also Table 1). Paradoxically, C3H/HeJ mice immunized with 10⁷ C3H.CE cells intravenously lose the capacity for vigorous proliferation to *Mls-1^a*, (refs 5, 6, see also Table 1). Thus, this system presents an *in vivo* model for the induction of non-responsiveness in mature T cells.

The frequencies of V β 6⁺ cells among CD4⁺ T cells within lymph node cells of treated as compared to untreated mice was increased nine days after injection and decreased after 12 days (Table 1). This result indicated that the reduction of *Mls*-specific responses in treated mice cannot be correlated with increase or decrease in the frequency of V β 6⁺CD4⁺ T cells and prompted us to do a more detailed analysis (Table 2). We found a considerable variation in the frequencies of V β 6⁺CD4⁺ T cells in treated, but not in untreated mice, dependant on the time elapsed between injection and assay. Again, a correlation between residual *Mls* reactivity and frequency of V β 6⁺CD4⁺ T cells was not found, indicating some kind of anergy in these cells. Similar results were obtained in a classical *Mls*-incompatible combination, BALB/c (*H-2^d*, *Mls-1^b*) versus DBA/2 (*H-2^d*, *Mls-1^a*) (Table 2, second part) and in B10.BR (*H-2^k*, *Mls-1^b*) versus AKR/J (*H-2^k*, *Mls-1^a*) (not shown). To investigate functional aspects of these anergic T cells, lymphocytes from both treated and untreated C3H mice were tested for their prolifera-

tive response to *Mls* and *H-2* antigens in a mixed lymphocyte reaction. On day 3, cultures from untreated, but not treated mice, stimulated against *Mls* contained about twice as many CD4⁺V β 6⁺ T cells than put in initially, the majority being of intermediate or large size (Table 3, Fig. 1c, d). *Mls*-stimulated cultures from both treated and untreated mice showed similar levels of IL-2 receptor expression in V β 6⁺ T cells (Fig. 1e, f), whereas IL-2 production was observed in cells from untreated mice only (Table 3). After 7 days of mixed lymphocyte culture, CD4⁺V β 6⁺ T cells from untreated mice had expanded sevenfold in response to *Mls*, whereas no expansion was observed in cells from treated mice. No differences were seen in the behaviour of CD4⁺V β 6⁻ T cells in both groups when stimulated with *H-2^d* (Table 3).

These data indicate the following. T cells with specificity for *Mls-1^a* but not non-*Mls* antigens, from mice pretreated with *Mls-1^a* antigen, are in a state of clonal anergy. The density of T-cell receptor and CD4 molecules on the cell surface is unchanged as compared to normal T cells (Fig. 1a, b), excluding down-regulation of receptor molecules as cause for low responsiveness. In response to antigen, the cells undergo limited blastogenesis and express IL-2 receptors, but they do not synthesize substantial amounts of DNA, nor expand in number or produce IL-2. Addition of IL-2-containing supernatant does not convert this state, and suppressing cells seem to play no role, as judged from cell mixing experiments (not shown). The fact, however,

TABLE 3 Blastogenesis, expansion and IL-2 production of CD4⁺V β 6⁺ and *Mls*-reactive T cells from C3H mice pretreated with *Mls-1^a* antigen

Day of injection	Recovery of subpopulations in % of input				CD4 ⁺ V β 6 ⁻ cells after stimulation				IL-2 production* upon stimulation with:	
	CD4 ⁺ V β 6 ⁺ cells after stimulation with C3H.CE (k; a)				CD4 ⁺ V β 6 ⁻ cells after stimulation with BALB/c (d; b)				C3H.CE (k; a)	BALB/c (d; b)
	small	3-day MLR intermed.	large	7-day MLR	small	3-day MLR intermed.	large	7-day MLR		
No injection	52.2	42.9	103.4	743	78.2	6.3	5.1	64	106,242	102,130
Day-12	40.2	33.2	23.4	104	73.8	5.0	6.4	59	1,381	57,790
	52.8	33.5	23.5	92	72.9	4.9	5.1	63	2,404	75,796

Lymph node cells from untreated or immunized C3H mice (as in Table 1) were stimulated in MLR against irradiated C3H.CE or BALB/c spleen cells. Three days later, responder cells were analysed by flow cytometry for expression of CD4 and V β 6 molecules and for forward scatter as a measure for volume. Numbers of small, intermediate, and large CD4⁺V β 6⁺ or CD4⁺V β 6⁻ cells, respectively, were determined (see also Fig. 1c, d) and are expressed as % of input as counted before culture. Numbers of CD4⁺V β 6⁺ and CD4⁺V β 6⁻ cells were also determined after 7 days of MLR and related to initial input numbers. On day 3 of MLR, an aliquot of supernatant was removed from all cultures and tested for IL-2 content. Each line represents data from a single mouse, corresponding to the three mice used in Expt. 2 of Table 1. Lymph node cells were cultured at 1.2 × 10⁶ per well of a 24-well culture plate together with 5 × 10⁵ irradiated stimulator cells. To test the IL-2 content of culture supernatant, 10⁴ IL-2 dependent CTL cells were incubated for 48 h with serial dilutions of supernatants. [³H]-thymidine uptake was determined as a measure of proliferation, and hence IL-2 content of probes. Data for a final dilution of 1:6 are indicated. Two colour fluorescence analysis was carried out as described in Fig. 1.

* c.p.m. [³H]-thymidine incorporation of CTL-2 cells (*H-2* haplotype; *Mls-1* allele).

that the anergic cells express IL-2 receptors in response to Mls, suggests the possibility that they may be able to act as suppressor cells by absorbing IL-2. Indeed, it has been shown that Lyt-1^+ T cells (most likely CD4^+) from Mls-1^b mice immunized with Mls-1^a antigen can behave as potent 'suppressor T cells', which are antigen nonspecific in their effector phase¹⁸.

Our data indicate that, at least in the model we have studied, the induction of unresponsiveness in adult animals (resulting in the clonal anergy of mature T cells) is mechanistically quite different from the establishment of natural self tolerance, which is achieved by clonal deletion^{1-4,10}. The clonal anergy described is also different from models of defective self tolerance, which

TABLE 2 Development of Mls specific response and relative number of $\text{V}\beta 6^+\text{CD4}^+$ T cells in lymph nodes of Mls-1^b mice immunized with Mls-1^a -expressing cells

Injection of Mls-1^a expressing cells (days before MLR)	Max. ratio of response to Mls versus H-2	Percentage of lymph node cells which are:		
		$\text{V}\beta 6^+$	$\text{V}\beta 6^+\text{CD4}^+$	$\text{V}\beta 6^+$ among CD4^+
C3H/HeJ mice				
No injection	3.2 ± 1.7 ($n=11$)	10.6 ± 0.9 ($n=18$)	6.5 ± 0.7 ($n=14$)	13.9 ± 1.3 ($n=14$)
-1	4.7	11.4	ND	ND
-1	5.6	11.2	ND	ND
-2	2.8	8.3	3.8	9.4
-2	ND	8.0	3.6	7.3
-2	1.6	8.0	2.7	8.2
-2	ND	9.6	3.5	7.0
-3	0.07	11.5	8.4	22.0
-3	0.05	11.4	5.8	19.4
-4	0.06	10.9	7.6	20.7
-4	0.15	11.7	8.4	22.1
-5	0.30	9.8	6.3	17.4
-5	ND	10.5	6.8	18.7
-5	ND	12.0	7.4	18.2
-6	0.09	7.5	3.3	7.2
-6	ND	10.6	7.1	16.0
-8	0.0	10.7	5.7	16.3
-9	0.26	12.7	8.7	20.6
-9	0.62	14.7	10.2	21.0
-9	ND	7.6	4.2	9.4
-9	ND	10.0	6.0	15.9
-12	0.30	6.6	4.1	5.9
-12	0.11	9.4	4.7	8.7
-16	0.19	9.5	4.2	13.2
-22	0.18	9.3	4.9	10.8
-22	0.85	9.7	5.8	12.4
-22	0.37	4.9	2.7	8.9
-25	0.21	10.8	6.9	14.8
-30	ND	8.9	4.4	8.3
-34	0.25	5.6	3.5	9.5
-34	0.5	7.6	4.4	11.8
-49	0.32	8.5	ND	ND
-57	0.41	9.9	ND	ND
BALB/c mice				
No injection	4.1 ± 2.0 ($n=12$)	8.2 ± 0.9 ($n=10$)	5.9 ± 1.0 ($n=8$)	11.8 ± 0.8 ($n=8$)
-2	2.3	7.2	ND	ND
-2	ND	6.1	3.0	5.2
-2	0.66	3.3	ND	ND
-2	ND	4.9	2.1	3.8
-3	1.1	6.1	4.1	10.6
-3	1.0	5.5	3.3	8.9
-4	0.27	4.4	3.2	10.4
-9	0.16	6.4	3.9	7.4
-9	0.19	6.6	4.1	7.9
-21	0.29	5.6	3.8	8.2
-21	1.25	4.6	3.0	6.8
-21	0.88	5.2	3.4	7.2
-39	0.13	4.5	ND	ND
-39	0.29	7.9	ND	ND
-54	1.0	5.8	4.1	7.9
-54	1.4	5.9	4.2	8.8
-54	0.63	5.1	3.2	7.3
-54	1.3	6.0	4.1	8.0
-54	1.2	6.4	3.5	7.4

See legend to Table 1. In addition, BALB/c (H-2^d , Mls-1^b) and DBA/2 (H-2^d , Mls-1^a) mice were used as a source of responder or donor/stimulator cells, respectively. Note that both the mice with increased frequency of $\text{V}\beta 6^+\text{CD4}^+$ T cells as well as those with unchanged frequencies show drastic reduction of Mls-specific proliferative responses. ND, not determined.

involve neither clonal deletion nor clonal anergy of the kind described here^{11,19-23}. In summary, we have described the induction *in vivo* of antigen specific nonresponsiveness in mature T lymphocytes, in the absence of clonal deletion. These observations appear to represent an *in vivo* correlate for the induction of anergy in T-cell lines *in vitro*¹¹⁻¹⁵ and are reminiscent of the

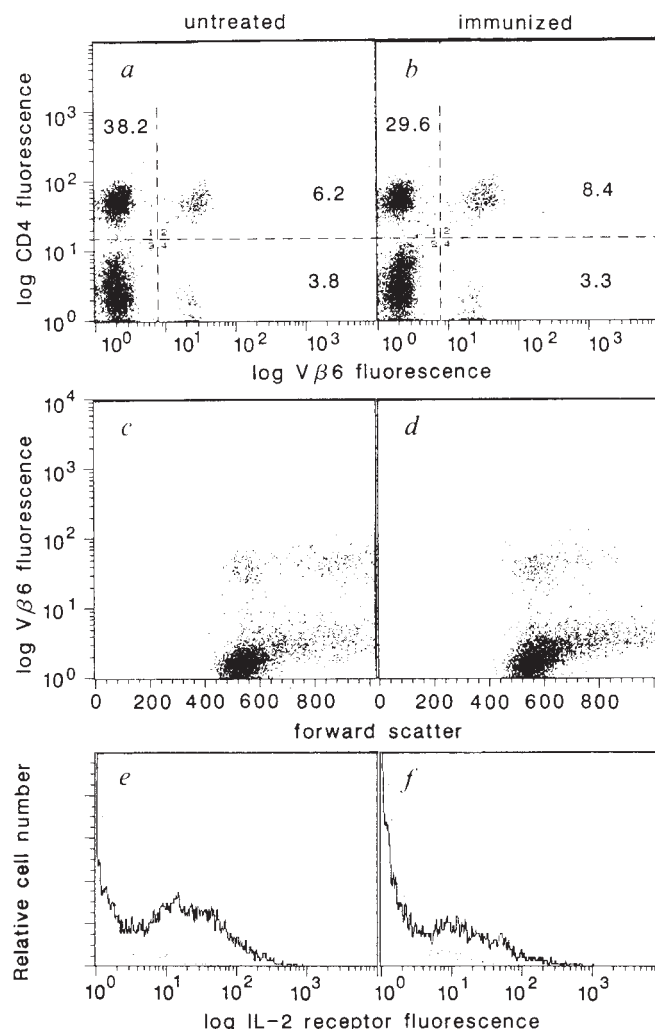


FIG. 1 Two-colour fluorescence analysis of responder T cells from immunized and untreated C3H mice. Lymph node cells from untreated (panels a, c, e) or Mls-1^a -immunized (b, d, f) C3H mice were analyzed immediately after isolation (a, b) or after a 3 day MLR against C3H.CE (c-f) stimulators. Data for cell culture are described in Table 3. Shown are plots of CD4 staining (phycoerythrin, red) versus $\text{V}\beta 6$ -staining (fluorescein, green) (a, b), or $\text{V}\beta 6$ -staining (green) versus forward light scatter (c, d) or histograms of IL-2 receptor-staining (solid line; control dotted) (e, f). $\text{V}\beta 6^-$ cells were gated out in panels e and f. Numbers in panels a and b indicate percentage of cells in the respective quadrants. The interval between immunization and mixed lymphocyte culture was 4 days (b, f), or 5 days (d).

METHODS. Cells were incubated at $1-5 \times 10^5$ cells per well in round bottom microtiter plates (Falcon) for 45 min at room temperature for each step, followed by washing. Medium was phosphate buffered saline containing 2% fetal calf serum. Successive staining steps were as follows: (1) anti $\text{V}\beta 6$ (ref. 27, supernatant of hybridoma 44-22-1, rat IgG); (2) Fluorescein-labelled F(ab)_2 fraction of goat anti-rat IgG (Cappel) absorbed on normal mouse serum; (3) biotinylated CD4 specific antibody (semipurified from ascites produced with GK1.5 hybridoma cells²⁸ and biotinylated according to standard procedures) or IL-2 receptor specific antibody 7D4 (ref. 29) or CD4-specific antibody RL 172.4 (ref. 30, rat IgM); (4) Streptavidin-conjugated phycoerythrin (Jackson) or Phycoerythrin-labelled F(ab)_2 fraction of goat anti-rat IgM, absorbed on mouse serum (Jackson). After the last wash, cells were resuspended in medium containing $25 \mu\text{g ml}^{-1}$ propidium iodide and analyzed in a FACScan analyzer (Becton Dickinson). Dead cells and cellular debris were gated out by using propidium iodide staining and forward light scatter. Calculations were carried out using FACScan software.

clonal anergy observed in B cells^{24,25}. It is likely that our findings will contribute to better understanding of the obviously complex mechanisms involved in the induction of unresponsiveness in the mature immune system. □

Received 14 February; accepted 14 April 1989.

- Kappler, J. W., Roehm, N. & Marrack, P. *Cell* **49**, 273–280 (1987).
- Kappler, J. W., Staerz, U., White, J. & Marrack, P. C. *Nature* **332**, 35–40 (1988).
- MacDonald, H. R. *et al. Nature* 40–45 (1988).
- Kisielow, P., Blüthmann, H., Staerz, U. D., Steinmetz, M. & von Boehmer, H. *Nature* **333**, 742–746 (1988).
- Jacobson, J., Lilliehöök, B. & Blomgren, H. *Scand. J. Immun.* **4**, 181–191 (1975).
- Lilliehöök, B., Jacobson, H. & Blomgren, H. *Scand. J. Immun.* **4**, 209–216 (1975).
- Miller, S. D., Sy, M.-S. & Claman, H. N. *Eur. J. Immunol.* **7**, 165–170 (1977).
- Brondz, B. D., Karanlov, A. V., Abronina, I. F. & Blandova, L. K. *Scand. J. Immun.* **13**, 517–534 (1981).
- Jenkins, M. K. & Schwartz, R. H. *J. exp. Med.* **165**, 302–319 (1987).
- von Boehmer, H., Teh, H. S. & Kisielow, P. *Immun. Today* **10**, 57–61 (1989).
- Markmann, J. *et al. Nature* **336**, 476–479 (1988).
- Lamb, J. R. *et al. J. exp. Med.* **157**, 1434–1447 (1983).
- Lamb, J. R. & Feldmann, M. *Nature* **308**, 72–74 (1984).
- Quill, H. & Schwartz, R. H. *J. Immun.* **138**, 3704–3712 (1987).
- Jenkins, M. K., Pardoll, D. M., Mizuguchi, J., Chused, T. M. & Schwartz, R. H. *Proc. natn. Acad. Sci. U.S.A.* **84**, 5409–5413 (1987).
- Festenstein, H. *Transplant. Rev.* **15**, 62–88 (1973).
- Abe, R. & Hodes, R. J. *Immun. Today* **9**, 230–235 (1988).
- Halle-Pannenko, O., Pritchard, L. L. & Bruley-Rosset, M. *Eur. J. Immun.* **17**, 1751–1755 (1987).
- von Boehmer, H. & Schubiger, K. *Eur. J. Immun.* **14**, 1048–1052 (1984).
- Ready, A. R., Jenkinson, E. J., Kingston, R. & Owen, J. J. I. *Nature* **310**, 231–233 (1984).
- Flajnik, M. F., DuPasquier, L. & Cohen, N. *Eur. J. Immun.* **15**, 540–547 (1985).
- Ohki, H., Martin, C., Corbel, C., Coltey, M. & Le Douarin, N. M. *Science* **237**, 1032–1035 (1987).
- Marrack, P. *et al. Cell* **53**, 627–634 (1988).
- Nossal, G. J. V. & Pike, B. L. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1602–1606 (1980).
- Goodnow, C. C. *et al. Nature* **334**, 676–682 (1988).
- Janeway, C. A., Jr., Fischer-Lindahl, K. & Hämmerling, U. *Immun. Today* **9**, 125–126 (1988).
- Payne, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 7695–7698 (1988).
- Dialynas, D. P. *et al. J. Immun.* **131**, 2445–2451 (1983).
- Malek, T. R., Robb, R. J. & Shevach, E. M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5694–5698 (1983).
- Ceredig, R., Lowenthal, J. W., Nabholz, M. & MacDonald, H. R. *Nature* **314**, 98–100 (1985).

ACKNOWLEDGEMENTS. We thank Jan Klein for his continuous support, Hans Hengartner for providing the 44-22.1 hybridoma, Stefan Faath for technical assistance, and Lynne Yakes for preparing the manuscript. This work was supported in part by Sonderforschungsbereich No. 120.

Lymphocytes bearing antigen-specific $\gamma\delta$ T-cell receptors accumulate in human infectious disease lesions

Robert. L. Modlin^{*†}, Claude Pirmez[†],
Florence M. Hofman[†], Victoria Torigian^{*†},
Koichi Uyemura[†], Thomas H. Rea^{*},
Barry R. Bloom[‡] & Michael B. Brenner[§]

^{*} Section of Dermatology, and [†] Department of Pathology, University of Southern California School of Medicine, Los Angeles, California 90033, USA

[‡] Department of Microbiology and Immunology, Albert Einstein College of Medicine, Bronx, New York 10461, USA

[§] Laboratory of Immunochimistry, Dana-Farber Cancer Institute and Department of Rheumatology and Immunology, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts 02115, USA

THE majority of T cells bear the T-cell receptor (TCR) $\alpha\beta$ complex which recognizes foreign antigen peptides only in the context of self major histocompatibility complex (MHC) molecules¹. Such T cells function in a variety of effector roles and secrete cytokines that mediate the activation and differentiation of other cells in the immune system. Recently, a small subpopulation T cells was found to bear a distinct TCR composed of γ and δ subunits². In man, TCR $\gamma\delta^+$ cells are distributed as ~5 per cent of the CD3⁺ cells in all organized lymphoid organs as well as in the skin- and gut-associated lymphoid tissues³. Although a limited number of germ-line genes encode the TCR γ and δ subunits, extensive functional variation particularly in the δ gene, results in unprecedented diversity for this receptor^{4,5}. The nature of the specificity

and immunological functions of these T cells remains enigmatic. We report here that in contrast to the normal low frequency of $\gamma\delta$ -bearing cells in lymphoid tissues, peripheral blood, or normal skin, the frequency is increased five to eightfold in particular granulomatous reactions of leprosy. TCR $\gamma\delta^+$ lymphocyte lines from these leprosy skin lesions proliferate *in vitro* specifically to mycobacterial antigens. This reactivity to foreign antigens appears to require presentation in the context of self-molecules. Moreover, culture supernatants from activated $\gamma\delta$ T lymphocytes induce adhesion and aggregation of bone-marrow monocytes in the presence of granulocyte monocyte-colony stimulating factor (CSF), suggesting that products of $\gamma\delta$ -bearing T cells may play a role in the immune response, possibly by stimulating granuloma formation.

Leprosy comprises a varied clinical spectrum that parallels cell-mediated immunity, the skin lesions of which are readily accessible for study⁶. Tuberculoid leprosy lesions contain relatively few organisms, and the host mounts a strong cell-mediated immune response which restricts growth of the bacilli. In lepromatous leprosy, however, numerous skin lesions are found containing high numbers of acid fast bacilli, and the host exhibits unresponsiveness to *Mycobacterium leprae* antigens⁷. Superimposed on this spectrum are reactional states including reversal reactions and lepromin skin tests. Reversal reactions appear to be acute delayed-type hypersensitivity reactions characterized by erythematous tumid lesions containing organized granulomas in which CD4⁺ T cells predominate⁸. Patients with tuberculoid leprosy and reversal reactions manifest positive lepromin skin tests (Mitsuda reactions) which are organized granulomatous reactions to *M. leprae* of three to four weeks duration⁹.

We compared the occurrence of lymphocytes bearing TCR $\gamma\delta$ or TCR $\alpha\beta$ in each of these immunological reactions to *M. leprae* using immunoperoxidase staining of skin biopsy specimens with specific monoclonal antibodies. TCR $\alpha\beta^+$ lymphocytes across the disease spectrum and in each of these reactional states were detected by staining with monoclonal antibody β F1 (ref. 10) (reacts with shared determinants present on all human TCR β chains), whereas TCR $\gamma\delta^+$ T cells were detected by staining with monoclonal antibody anti-TCR δ 1 (ref. 11) (reacts with shared determinants present on all human TCR δ chains). In both lepromin skin tests and in reversal reactions, TCR δ 1⁺ lymphocytes comprised 25–35 per cent of CD3⁺ cells in the lesions compared to ~5 per cent of the CD3⁺ cells in lesions of other forms of the spectrum (Figs 1 and 2) or in normal lymphoid organs. This suggests that TCR $\gamma\delta$ lymphocytes may be involved in early active granulomatous responses (such as lepromin skin tests and reversal reactions), in contrast to the more chronic and/or immunologically unresponsive leprosy lesions.

Localized American cutaneous leishmaniasis is another infectious disease with granulomatous skin reactions¹². We found similarly that TCR $\gamma\delta^+$ cells also were present in large numbers in these lesions. Twenty per cent of the CD3⁺ cells in such lesions stained with anti-TCR δ 1, corresponding to a sevenfold increase over the percentage of these lymphocytes in peripheral blood from the same patients, or in Montenegro skin reactions (48 h delayed type hypersensitivity responses to leishmania antigen) or in mucocutaneous leishmaniasis (a chronic destructive form of the disease) where typically only 5% of the CD3⁺ cells express a TCR δ chain (Figs 1 and 2).

The specific recognition of antigens by TCR $\gamma\delta$ lymphocytes has been difficult to demonstrate. These lymphocytes have been noted to lyse tumour cells^{13,14} as well as allogeneic target cells¹⁵. But evidence for $\gamma\delta$ T cells with specificity for soluble antigens or exhibiting self-restricted recognition of foreign antigens has been lacking. To examine the antigen specificity of lymphocytes accumulated in leprosy lesions, we derived T-cell lines from skin lesions and peripheral blood as previously described (Fig. 3; see ref. 16). One T-cell line was obtained from the biopsy of a lepromin skin test of a patient with tuberculoid leprosy in

reversal reaction. This line was expanded *in vitro* with *M. leprae* antigens and with interleukin-2 (IL-2), and then was subjected to cell sorting to isolate lymphocytes unreactive with monoclonal antibodies against TCR $\alpha\beta$ (WT31) and CD4 (which is rarely expressed on TCR $\gamma\delta^+$ cells). The resulting cell line was enriched for $\gamma\delta$ T cells (99% TCR $\delta 1^+$ and 1% WT31 $^+$). This line proliferates specifically in response to *M. leprae* cell-wall antigens (15,256 c.p.m.) and tuberculin purified protein derivative (PPD) (27,478 c.p.m.) but not to the 65K (relative molecular mass) (967 c.p.m.) or 18K (1,253 c.p.m.) recombinant heat-shock proteins of mycobacteria, to tetanus toxoid (848 c.p.m.), or to medium without specific antigen (494 c.p.m.) (Fig. 3a shows typical results from one of eight experiments). Similar experiments were carried out on T-cell lines and clones expressing TCR $\gamma\delta$ that were derived from normal donors using tetanus toxoid or phytohaemagglutinin (PHA), and these lines showed no response to *M. leprae* cell wall or PPD (data not shown). A second T-cell line (99% TCR $\delta 1^+$ and 1% WT31 $^+$) was derived from the peripheral blood of a leprosy patient. Like the TCR $\gamma\delta$ line established from the reversal reaction skin lesion, this line also proliferated specifically both to *M. leprae* cell-wall antigens and to PPD. In one of four typical experiments the responses to *M. leprae* cell wall (13,983 c.p.m.) and PPD (7,052 c.p.m.)

were unequivocally higher than control stimulations using the same culture conditions except for medium lacking in these antigens (256 c.p.m.).

Because the TCR $\alpha\beta$ complex recognizes antigens in the context of self-MHC molecules, it was of interest to ask whether antigen recognition by TCR $\gamma\delta$ cells requires self antigen presentation. Experiments using antigen-presenting cells (APCs) of differing HLA types revealed that the TCR $\gamma\delta$ lymphocytes appear to require self or matched antigen-presenting cells. For example, the reversal reaction-derived $\gamma\delta$ T-cell line responded to PPD maximally only in the presence of autologous APCs or APCs from one donor matched at HLA-B7, DRw15 and DQ1 (this donor's cells were used in the *in vitro* propagation of the line). In contrast APCs derived from ten additional donors yielded only weak responses, including individuals matched only at B7, two matched at DRw15, DQ1 and one matched at B7, DRw15 and DQ1 (Fig. 3b). Similarly, the peripheral blood-derived *M. leprae* and PPD-reactive cell line responded in the presence of autologous APCs, whereas no responses were elicited using three donors matched at one class II allele or one donor who was histoincompatible at all typed HLA alleles (Fig. 3c). Thus, the response to antigen appeared to be influenced by genetic factors; however, the limited analysis performed is not

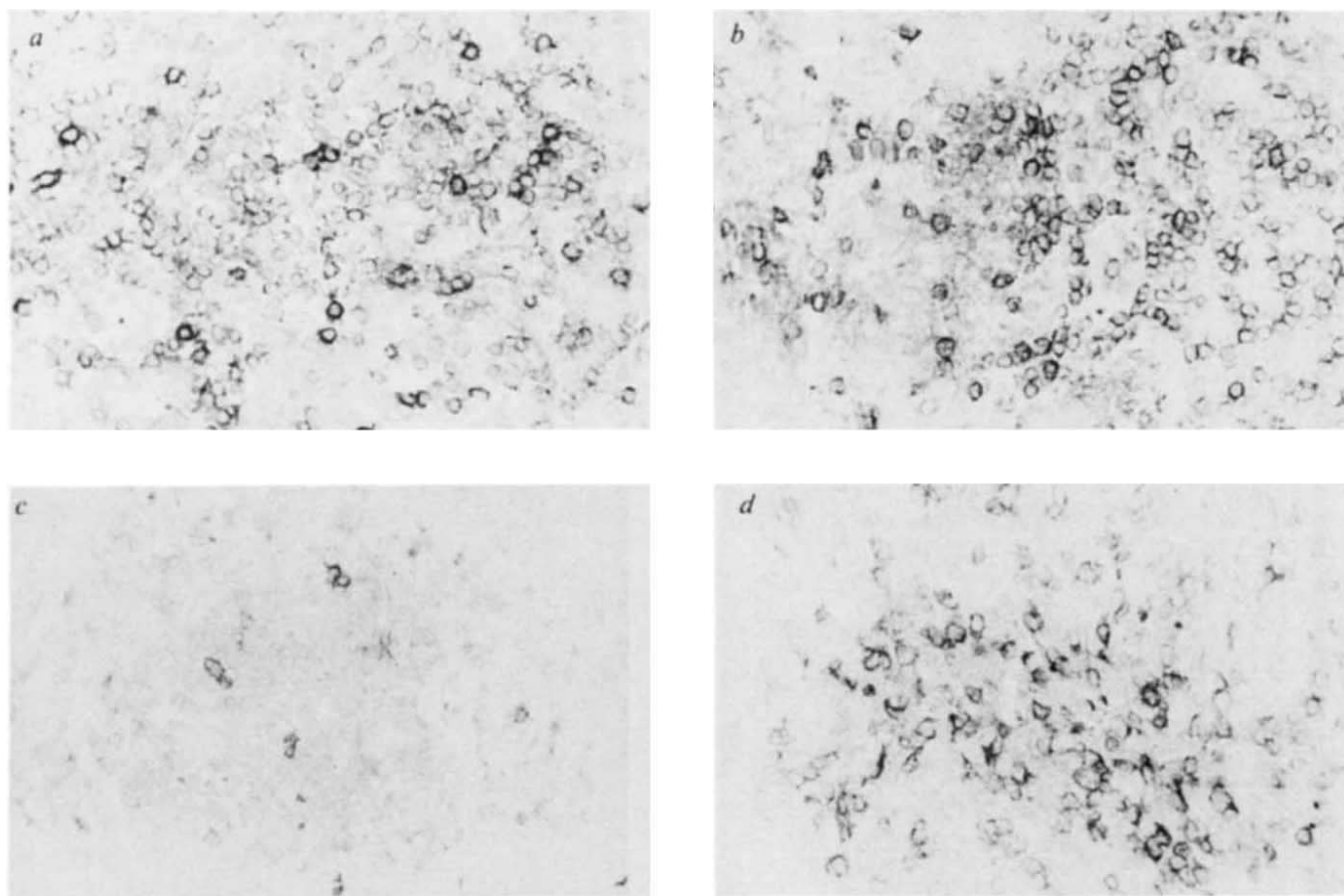


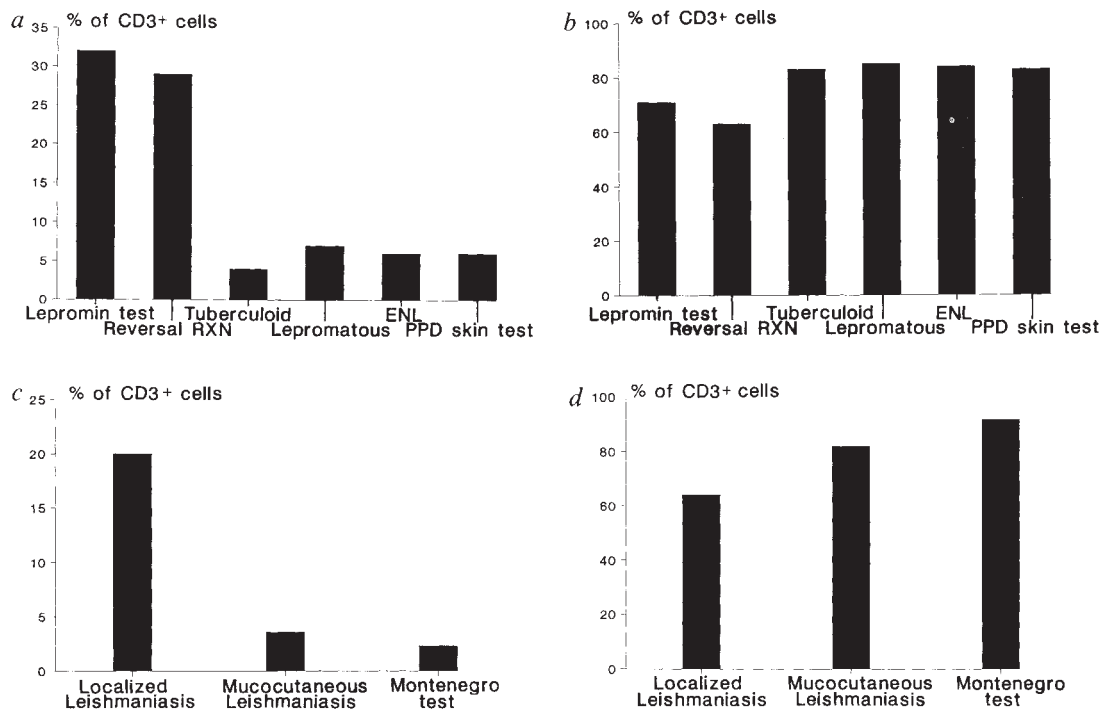
FIG. 1 Immunostaining of TCR-bearing cells in lesions. *a*, TCR $\gamma\delta$ lymphocytes in a lepromin skin test identified with monoclonal antibody anti-TCR $\delta 1$. *b*, TCR $\alpha\beta$ lymphocytes in the same lepromin skin test identified with monoclonal antibody $\beta F1$. *c*, TCR $\gamma\delta$ lymphocytes in a Montenegro reaction identified with anti-TCR $\delta 1$. *d*, TCR $\alpha\beta$ lymphocytes in the same Montenegro reaction identified with $\beta F1$.

METHODS. Patients with leprosy were classified according to the clinicopathological criteria of Ridley and Jopling²⁴ at the Los Angeles County, University of Southern California Hansen's Disease Clinic. Patients with leishmaniasis (*Leishmania braziliensis*) were diagnosed at the Oswaldo Cruz Institute, Rio de Janeiro according to clinical, parasitological and/or immunological criteria. Lepromin skin tests (Mitsuda reactions) were biopsied at 3

weeks, Montenegro tests at 48 hours. Skin biopsy specimens, obtained by punch or scalpel technique were divided for histological diagnosis on conventional paraffin sections and for immunoperoxidase study by snap freezing in liquid nitrogen²⁵. Mouse monoclonal antibodies were used at concentrations predetermined by titrations on cells from reactive tonsils. $\beta F1$ (ref. 10) (1:40) reacts with TCR $\alpha\beta$ lymphocytes and anti-TCR $\delta 1$ (ref. 11) (1:20) reacts with TCR $\gamma\delta$ cells. Each specimen was also stained with a panel of monoclonal antibodies including anti-CD3 (Leu4, 1:200, Becton Dickinson). Single immunoperoxidase staining: a two-step indirect immunoperoxidase technique was used as previously described²⁵. The percentage of cells expressing each TCR was estimated by dividing the number of cells reacting with either $\beta F1$ or anti-TCR $\delta 1$ by the total number of CD3 $^+$ cells.

FIG. 2 Percentages of TCR-bearing CD3⁺ cells in lesions. *a*, Percentage of CD3⁺ cells expressing TCR δ in leprosy lesions. *b*, Percentage of CD3⁺ cells expressing TCR β in leprosy lesions. *c*, Percentage of CD3⁺ cells expressing TCR δ in leishmaniasis lesions. *d*, Percentage of CD3⁺ cells expressing TCR β in leishmaniasis lesions.

METHODS. Patient lesions studied by immunoperoxidase included tuberculoid leprosy ($n=7$), lepromatous leprosy ($n=6$), lepromin skin tests ($n=4$), reversal reactions ($n=6$), PPD skin tests ($n=3$), erythema nodosum leprosum ($n=4$), localized American cutaneous leishmaniasis ($n=7$), mucocutaneous leishmaniasis ($n=2$) and Montenegro reactions ($n=5$). Standard deviation of percentages ranged from 1 to 10. It was apparent that in some lesions, neither TCR β nor TCR δ could be detected on a small fraction of CD3⁺ cells. Peripheral blood staining: peripheral blood mononuclear cells from leprosy and leishmaniasis patients were stained simultaneously with fluorescein-conjugated WT31 (Becton Dickinson) and phycoerythrin-conjugated anti-CD3 (Coulter)



or anti-TCR δ 1 followed by fluorescein-conjugated goat anti-mouse (Tago) and phycoerythrin-conjugated anti-CD3. Cells were analysed by two colour flow cytometry. Fewer than 10% of peripheral blood CD3 cells were found to express TCR $\gamma\delta$ in 20 leprosy patients and six leishmaniasis patients as well as 10 normal subjects.

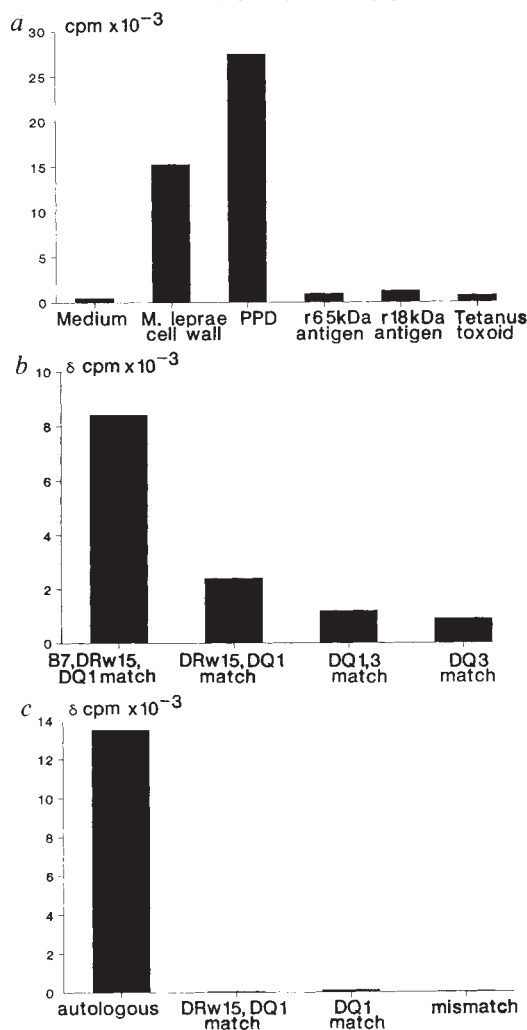


FIG. 3 Reactivity and genetic restriction of antigen-reactive TCR $\gamma\delta$ lymphocytes. *a*, Reactivity of TCR $\gamma\delta$ T-cell line from a lepromin skin test to mycobacterial antigens. *b*, Genetic restriction of TCR $\gamma\delta$ cell line PPD response from a lepromin skin test. The HLA type in experiments where donors were matched is shown. The donor shown as matched for HLA alleles B7, DRw15 and DQ1 was used for *in vitro* propagation of the cell line. *c*, Genetic restriction of an *M. leprae* response from a second TCR $\gamma\delta$ cell line from the peripheral blood of a leprosy patient.

METHODS. Extraction of lymphocytes from biopsy specimens was performed as previously reported^{26,27}. Because insufficient numbers of lymphocytes from lesions were obtained for direct functional studies, T-cell lines were established by plating 10,000 lymphocytes per 200 μ l in flat bottom wells (Linbro, Flow Laboratories) with 5×10^5 lethally irradiated (3,000 rad) HLA-matched peripheral blood mononuclear cells as feeders in RPMI 1640 containing 10% heat-inactivated pooled human AB serum, 10% interleukin-2 (Electronucleonics) and *M. leprae* cell walls (10μ g ml⁻¹, prepared by Dr Patrick Brennan, Fort Collins, Colorado). Patient-autologous APCs were not routinely used because some of these patients have circulating suppressor cells¹⁷; however, this line proliferated in response to *M. leprae* cell wall antigens (8,714 cpm) in the presence of autologous adherent cells. The patient's HLA type was A11, B7, 38, DRw15, 9, DQ1, 3. A normal donor of HLA-type A24, 26; B7, 44; DRw15, 3; DQ1, 2 was used as APCs for *in vitro* cultivation of the leprosy skin lesion-derived T-cell line. A second T-cell line from the peripheral blood of a tuberculoid patient was derived using autologous APCs. After expansion, lines were depleted of WT31⁺ and CD4⁺ cells by flow cytometry yielding T-cell lines of >99% purity as assessed by staining with anti-TCR δ 1. To assure assay of a homogeneous TCR δ 1⁺ population, the lines were tested directly after cell sorting (>99% TCR δ 1⁺) (in some cases assays were repeated after several months in culture when the phenotype was CD3=95%, TCR δ 1=97%, WT31=4%, CD4=4%, and CD8=52%). For assay of T-cell proliferation, arithmetic means of tritiated thymidine incorporation of triplicate cultures are shown after stimulation with *M. leprae*, PPD, r65K antigen of *M. bovis*²⁸ and r18K antigen of *M. leprae*²⁹ as previously described²⁷. Control stimulations used were tetanus toxoid, PHA, or 10% IL-2 to demonstrate antigen specificity and cell viability.

adequate to define clearly a putative self-restricting element(s).

Granuloma formation in the tuberculoid form of leprosy, as well as in leishmaniasis and tuberculosis is presumed to be important for restricting the spread or growth of the pathogens¹⁷. The finding that antigen-specific TCR $\gamma\delta$ cells accumulate in large numbers in immunological reactions characterized by active granuloma formation suggested the possibility that these cells or their secreted products may play a role in the development of such lesions¹⁸. Bone marrow-derived macrophages were cultured in presence of CSF, to support growth (ref. 19), and supernatants obtained from anti-CD3 or antigen-stimulated TCR $\gamma\delta$ lymphocyte lines derived from leishmaniasis lesions. Striking macrophage aggregation and enhanced cell division were seen in cultures containing CSF with supernatants of activated TCR $\gamma\delta$ cells, but not those containing supernatants of non-stimulated TCR $\gamma\delta$ cells or either component alone (Fig. 4).

Our results indicate that TCR $\gamma\delta$ lymphocytes: (1) accumulate in reactive granulomatous lesions of leprosy and local cutaneous leishmaniasis; (2) when cultured from leprosy lesions, can proliferate to mycobacterial antigens *in vitro*; (3) appear to be genetically restricted and requiring self antigen presentation; and (4) upon stimulation with antigen, secrete lymphokines that cause macrophage adhesion, aggregation and proliferation.

Yet these studies raise a number of questions. While it is clear that TCR $\gamma\delta$ have the potential for recognition of an enormous diversity of antigenic specificities, and the data presented here indicate that mycobacterial antigens are recognized *in vitro*, it

remains to be learned whether a large number of epitopes on conventional antigens are each recognized by a limited number of specific $\gamma\delta$ T cell clones, whether the response is restricted to highly conserved antigens such as 'stress proteins', and whether mycobacterial antigens act as 'superantigens'²⁰ that stimulate whole subsets of $\gamma\delta$ T cells.

Recent evidence suggests that mycobacterial antigens may stimulate one subset of murine $\gamma\delta$ TT-hybridomas (perhaps somewhat analogous to superantigen responses)²¹ or primary murine T cells in an MHC class II independent fashion²². Our data provide evidence suggesting genetic restriction of antigen presentation to TCR $\gamma\delta$ lines obtained from different individuals, although the restricting element could not be clearly associated with the MHC and it is unclear whether specialized APCs are required for expanding antigen-specific TCR $\gamma\delta$ cells.

It is intriguing that TCR $\gamma\delta$ cells were found at sites of leprosy reversal reactions and lepromin skin tests, but not established tuberculoid lesions, suggesting that the cells could be involved in active granuloma formation. Since in previous studies both IL-4²³ and tumour necrosis factor¹⁷ have been implicated in granuloma formation, it will be important to define the lymphokine(s) produced by activated TCR $\gamma\delta$ cells responsible for inducing macrophage aggregation *in vitro*.

Overall, the present results suggest that at least some TCR $\gamma\delta$ cells with specificity for foreign antigens may have a significant role in responses that are biologically important in both resistance to certain infectious pathogens and in tissue damage found in these and possibly in autoimmune diseases. □

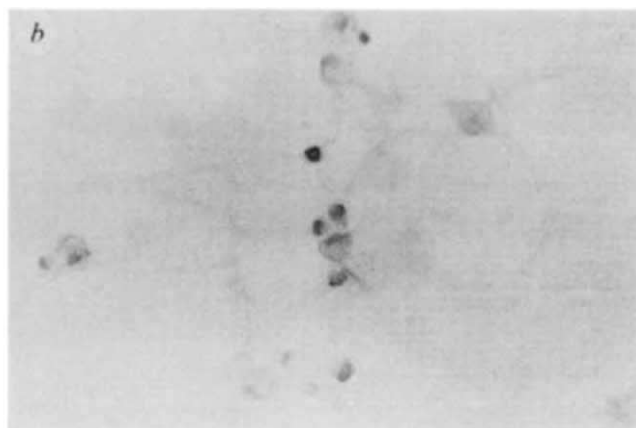
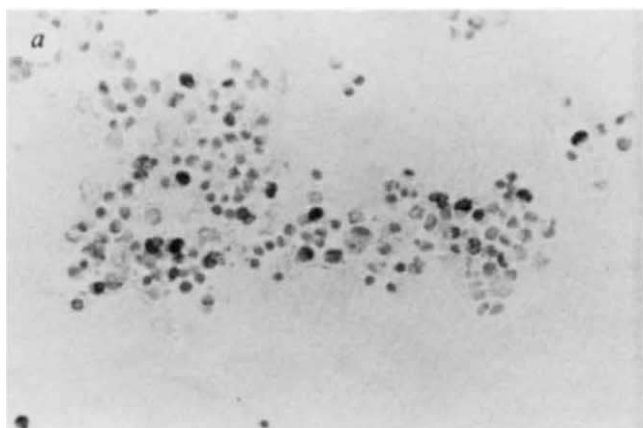
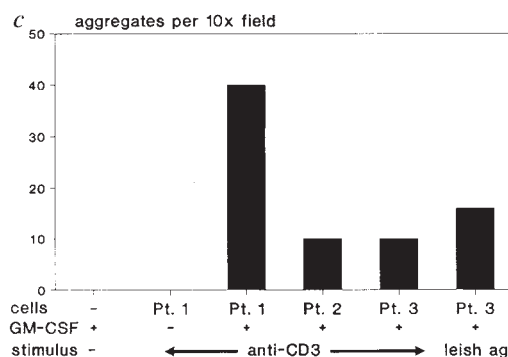


FIG. 4 Aggregation and proliferation of human bone marrow macrophages by supernatants of stimulated TCR $\gamma\delta$ lymphocytes derived from leishmaniasis lesions. *a*, Photomicrographs of macrophage aggregation and proliferation induced by CD3-stimulated TCR $\gamma\delta$ T-cell supernatants in the presence of CSF. The dark cells have incorporated BrdU (anti-BrdU immunoperoxidase counterstained with haematoxylin, $\times 40$ objective). *b*, Photomicrographs of macrophage aggregation and proliferation induced by unactivated TCR $\gamma\delta$ T-cell supernatants in the presence of CSF. No aggregation is seen in cultures treated with CSF alone. *c*, Number of macrophage aggregates per $\times 10$ objective field induced by anti-CD3 or *L. b. braziliensis*-stimulated TCR $\gamma\delta$ cell supernatants in the presence or absence of CSF. METHODS. TCR $\gamma\delta$ cell lines were derived from localized American cutaneous leishmaniasis lesions in the presence of *L. b. braziliensis* (patients 1, 2 and 3) and enriched to $>99\%$ purity by anti-TCR $\delta 1$ staining. Supernatants were obtained from these lines either by stimulating with anti-CD3 mAb (64.1) without APCs or stimulating with antigen in the presence of APCs. Controls consisted of unstimulated cell-line supernatants as well as supernatants from antigen-stimulated APCs without the T-cell line. Supernatants were assayed for their effect on bone marrow cultures derived from bone marrow aspirates as described previously³⁰. Bone marrow mononuclear cells were distributed into Lab-Tek chamber wells (Springfield, Illinois) at 10^5 cells per well. The wells contained either 33% supernatant plus CSF (150U, GM-CSF, Genzyme) supernatant alone or CSF alone in 33 μ l RPMI medium containing 10% AB serum¹⁹. Between 3 and 6 days of culture, large aggregates of 10–40 cells were apparent in wells containing supernatants from anti-CD3 or antigen-stimulated TCR $\gamma\delta$ lymphocytes plus



CSF. To identify the cellular composition of the clusters, the bone marrow cultures were stained with a macrophage-specific monoclonal antibody (Leu M5 at 1:50, Becton Dickinson) and contained $>80\%$ macrophages. Proliferation was determined using the cell proliferation kit (Amersham) which uses anti-BrdU antibodies with immunoperoxidase to visualize BrdU incorporation into dividing cells. The percentage of positive cells was determined by counting the number of positive cells per 250 cells total. Anti-CD3 stimulated supernatants plus CSF demonstrated 20% proliferation, whereas cultures with supernatant alone (10%) and CSF alone (6%) showed lower rates of proliferation.

Received 10 April; accepted 5 May 1989.

1. Marrack, P. & Kappler, J. *Adv. Immun.* **38**, 1–30 (1986).
2. Brenner, M. B. *et al. Nature* **322**, 145–149 (1986).
3. Groh, V. *et al. J. exp. Med.* **169**, 1277–1294 (1989).
4. Chien, Y. *et al. Nature* **330**, 722–727 (1987).
5. Hata, S. *et al. Science* **240**, 1541–1544 (1988).
6. Modlin, R. L. & Rea, T. H. *Springer Semin. Immunopathol.* **10**, 359–374 (1988).
7. Mehra, V., Mason, L. H., Fields, J. P. & Bloom, B. R. *J. Immun.* **123**, 1813–1817 (1979).
8. Cooper, C. L. *et al. J. exp. Med.* (in the press).
9. Dugan, E., Modlin, R. L. & Rea, T. H. *Int. J. Leprosy* **53**, 404–409 (1985).
10. Brenner, M. B. *et al. J. Immun.* **138**, 1502–1509 (1987).
11. Band, H. *et al. Science* **238**, 682–684 (1987).
12. Convit, J. *Etiop. Med.* **12**, 187–195 (1974).
13. Borst, J. *et al. Nature* **325**, 683–688 (1987).
14. Brenner, M. B. *et al. Nature* **325**, 689–694 (1987).
15. Matis, L. A., Cron, R. & Bluestone, J. A. *Nature* **330**, 262–264 (1987).
16. Modlin, R. L. *et al. Nature* **329**, 541–545 (1987).
17. Kindler, V., Sappino, A.-P., Grau, G. E., Pigue, P.-F. & Vassalli, P. *Cell* **56**, 731–740 (1989).
18. Bennett, B. & Bloom, B. R. *Proc. natn. Acad. Sci. U.S.A.* **59**, 756–762 (1968).
19. Gentleman, H. E. *et al. J. exp. Med.* **167**, 1428–1441 (1988).
20. White, J. *et al. Cell* **56**, 27–35 (1989).
21. O'Brien, R. L. *et al. Cell* (in the press).
22. Janis, E. M., Kaufman, S. H. E., Schwartz, R. H. & Pardoll, D. M. *Science* (in the press).
23. McInnes, A. & Rennick, D. M. *J. exp. Med.* **167**, 598–611 (1988).
24. Ridley, D. S. & Jopling, W. H. *Int. J. Lepr.* **34**, 255–273 (1986).
25. Modlin, R. L. *et al. J. Immun.* **132**, 3085–3090 (1984).
26. Modlin, R. L. *et al. J. Immun.* **137**, 2831–2834 (1986).
27. Modlin, R. L. *et al. Nature* **322**, 459–461 (1986).
28. Thole, J. E. *et al. Infect. Immun.* **56**, 1633–1640 (1988).
29. Booth, R. J., Harris, D. P., Love, J. M. & Waston, J. D. *J. Immun.* **140**, 597–601 (1988).
30. Hofman, F. M., Brock, M., Taylor, C. R. & Lyons, B. *J. Immun.* **141**, 1185–1190 (1988).

ACKNOWLEDGEMENTS. We thank Drs Craig Morita and Steve Porcelli for providing control $\gamma\delta$ T cell lines that were not antigen reactive and Drs J. van Embden and J. Watson for kindly providing recombinant mycobacterial antigens. Supported by the NIH, the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (IMMTEP) and Vaccine Development (IMMTR), the National Hansen's Disease Center, Knights of Saint Lazarus of Jerusalem and Crown Foundation.

Identification of human CD4 residues affecting class II MHC versus HIV-1 gp120 binding

Linda K. Clayton^{*,†}, Monica Sieh^{*}, Donald A. Pious[‡] & Ellis L. Reinherz^{*§}

^{*}Laboratory of Immunobiology, Dana-Farber Cancer Institute and Departments of [†]Pathology and [§]Medicine, Harvard Medical School, 44 Binney Street, Boston, Massachusetts 02115, USA

[‡]Department of Pediatrics, University of Washington, Seattle, Washington 98195, USA

INTERACTIONS of CD4 with the class II major histocompatibility complex (MHC) are crucial during thymic ontogeny¹ and subsequently for helper and cytotoxic functions of CD4⁺CD8[−] T lymphocytes^{2–5}. CD4 is the receptor for the T-lymphotropic human immunodeficiency virus and binds its envelope glycoprotein, gp120^{6–10}. The residues involved in gp120 binding have been localized to a region within the immunoglobulin-like domain I of CD4, which corresponds to CDR2 of an immunoglobulin variable region^{11–13}, but the CD4 residues important in MHC class II interaction have not been characterized. Here, using a cell-binding assay dependent specifically on the CD4–MHC class II association, we analyse the effects of mutations in CD4 on class II versus gp120 binding. Mutations in CDR2 that destroy gp120 binding affect CD4–MHC class II binding similarly. In addition, binding of soluble gp120 to CD4-transfected cells abrogates their ability to interact with class II-bearing B lymphocytes. In contrast, other mutations within domains I or II that have no effect on gp120 binding eliminate or substantially decrease class II interaction. Thus, the CD4 binding site for class II MHC is more complex than the gp120 binding site, possibly reflecting a broader area of contact with the former ligand and a requirement for appropriate juxtaposition of the two N-terminal domains. The ability of gp120 to inhibit the binding of class II MHC to CD4 could be important in disrupting normal T-cell physiology, acting both to inhibit immune responses and to prevent differentiation of CD4⁺CD8[−] thymocytes into CD4⁺CD8[−] T lymphocytes.

To establish the structural basis of the CD4-class II MHC interaction, which could facilitate cell–cell contact^{2–5,14} and influence T-cell signal transduction^{15–17}, we used a cell conjugate assay specific for binding of CD4 to class II MHC and investigated the effects of mutations in the first two immunoglobulin-like domains of the CD4 molecule on MHC binding. Wild-type or mutant CD4 molecules are expressed by transfection into COS-1 cells and binding to MHC class II is measured by examining adhesion of class II MHC-expressing B cells to the COS-1 cells. As shown in Fig. 1a (left panel), the Epstein-Barr virus transformed B-lymphoblastoid cell line, T51 (DR1, 3; DQ1, 2; DPX, 4)¹⁸, binds readily to CD4-transfected COS-1 cells, but there is no T51 B-cell binding to COS-1 cells transfected with the CDM8 vector alone (Fig. 1a, right panel).

The surface expression and structural integrity of each mutant CD4 was confirmed by fluorescence-activated cell sorter (FACS) analysis using the monoclonal antibody OKT4. Each mutant was also tested for gp120 binding. Examples are shown in Fig. 1b, and results for all mutants are presented in Table 1. Wild-type CD4- and mutant M1B-transfected COS-1 cells bind equivalent amounts of OKT4 and gp120 (Fig. 1b); M3, although reactive with OKT4, fails to bind gp120¹³. Although M3 reactivity with monoclonal antibodies OKT4 and 19Thy5D7 is slightly lower than that of CD4, we found identical binding activity with the anti-CD4 monoclonal antibody MT321, which binds an epitope in domain IV (ref. 8; data not shown). These results show that equivalent amounts of surface CD4 are obtained when COS-1 cells are transfected with M3 and CD4 complementary DNAs, indicating that the M3 mutation may directly or indirectly perturb the epitopes seen by OKT4 and 19Thy5D7.

To test the specificity of our assay further, we investigated the effects of monoclonal antibodies directed against the individual components of the system. As shown in Fig. 2a, binding of the T51 B cells is markedly inhibited by monoclonal antibodies to CD4 (OKT4) and class II MHC (9/49), but is unaffected by monoclonal antibodies directed against CD8 (7Pt3F9) or class I MHC (W6/32). We obtained identical results for the genotypically unrelated B cell line JY (data not shown). In addition, the MHC class II antigen loss B-cell mutant line, 6.1.6, which fails to express any class II MHC alleles from either haplotype, does not bind to CD4-transfected COS-1 cells (Fig. 2b). In contrast, the mutant cell lines that express at least one complete haplotype, 9.28.6 (DR1, –; DQ1, –; DPX, –) and 8.1.6 (DR3, –; DQ2, –; DP4, X), bind, but at reduced levels. This result shows that different polymorphic alleles of class II MHC can bind to CD4. It is notable that the loss variants 4.36.4 and 11.11.4, which express only DQ1, DPX and DQ1 respectively, do not bind to CD4-expressing COS-1 cells. Whether the lack of binding of DR-negative B-cell variants is due to the low level of expression of DP and/or DQ in these cells as compared with DR (our unpublished observation), or is indicative of a lower affinity of CD4 for DP or DQ relative to DR cannot be determined at present. It has been shown, however, that these DP and DQ alleles are functional as restriction elements in T-cell responses¹⁸. Furthermore, in the murine system, liposomes containing either murine I-A or I-E molecules bind CD4-expressing cells¹⁹.

The murine CD4 sequence is 50% identical with its human homologue²⁰ and binds poorly to T51 B cells, the extent being similar to the binding shown in Fig. 1a (middle panel) for mutant M11 (see below). Given the obvious difference in binding of human class II-expressing B cells to human versus mouse CD4-transfected COS-1 cells, we used oligonucleotide-directed mutagenesis to create 17 individual CD4 mutants incorporating all non-conservative (polarity and/or charge change) murine for human substitutions between amino-acid residues 17 and 167. Each human CD4 mutant contains between one and four murine amino-acid substitutions. The positions of these are shown in Fig. 3, with specific residue changes listed in Table 1. Each mutant was transfected into COS-1 cells, assayed for

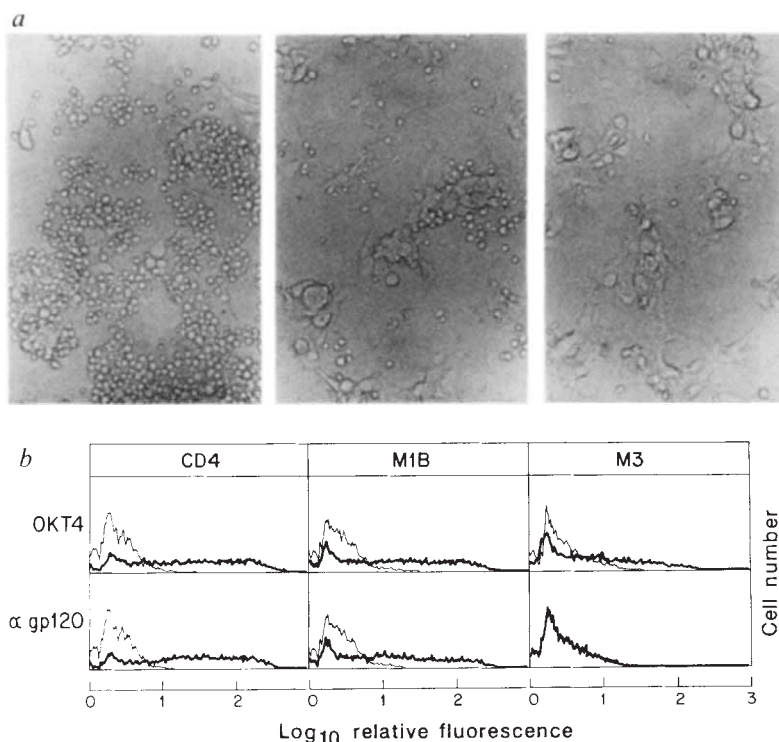
FIG. 1 Expression of CD4 and cellular adhesion with MHC class II-expressing B cells in transfected COS-1 cells. *a*, Binding of T51 B cells to COS-1 cells transfected with CD4 (left panel), CD4 mutant M11 (middle panel) or CDM8 vector only (right panel). *b*, FACS analysis of anti-CD4 monoclonal antibody OKT4 and gp120 binding to COS-1 cells transfected with wild-type CD4, M1B and M3 mutants.

METHODS. Plasmid pSP65-T4 (gift from Dan Littman) was digested with *Bam*HI and *Xho*I to release the CD4 insert. The insert was blunted with the Klenow fragment of DNA polymerase I, ligated to *Xba*I linkers (New England Biolabs) and subcloned into the *Xba*I site of the vector CDM8 (ref. 33). COS-1 cells were transfected with CDM8 constructs as described¹³, but with the following modifications: 0.3×10^6 cells were plated into each well of Falcon 6-well dishes and 6 μ g plasmid DNA (CsCl-banded twice) in 0.35 ml RPMI plus 0.35 ml RPMI-800 μ g ml⁻¹ DEAE dextran was used for transfection. Binding of B cells to transfected COS-1 cells was assayed two days after transfection as described¹⁴, with the following modifications: transfected COS-1 cells were washed once with RPMI and $1-2 \times 10^7$ B cells were added to each 35 mm well in 0.8 ml RPMI, 2% fetal calf serum, 1% glutamine, 1% penicillin-streptomycin, 10 μ g ml⁻¹ gentamicin (final medium). The mixed cells were incubated for one hour at 37 °C, the B cells aspirated and the well was washed 3-5 times by dropping 2 ml final medium into the well. Binding was scored as +, +/- or - after viewing each well by phase contrast microscopy at a magnification of 100 $\times$. A plus value represents binding equivalent to that obtained with COS-1 cells transfected with wild-type CD4 DNA. A minus value is indistinguishable from transfection results obtained with CDM8 alone. A +/- value represents substantially reduced, but still detectable, binding (10-20% wild-type binding). A negative control of CDM8-transfected COS-1 cells and a positive control of wild-type CD4-transfected COS-1 cells were included in every assay. We consistently obtained wild-type CD4 expression in 50-60% of COS-1 cells at equivalent levels of mean channel fluorescence in >36 separate transfections. Every B cell-binding assay for wild-type and mutant CD4s was accompanied by FACS analysis and always gave reproducible results ($n \geq 3$ for each mutant). For FACS analysis, COS-1 cells transfected with the mutant indicated were scraped from the well, stained with monoclonal antibodies OKT4, 19Thy5D7 or MT321 (gift from P. Rieber) (1:500 and 1:100 dilution of ascites for

surface expression of CD4 and for gp120 binding by FACS analysis, and tested for adhesion of T51 B cells. Results are given in Table 1. Seven mutants dramatically reduced or eliminated class II MHC binding. These included three mutations in the CDR1 homologous region of domain I (ref. 13), mutants M1.1, M1.2 and M1B; one mutation in the CDR2 equivalent region¹³, mutant M3; one mutation unrelated to any CDR, M5; one mutation immediately distal to the CDR3 homologue, M8; and one mutation in domain II, M11. An example of a mutation that reduces B-cell binding is shown in Fig. 1*a* (middle panel)

FIG. 2 Specificity of binding of CD4 transfected COS-1 cells to MHC class II-expressing B cells. *a*, Inhibition of T51 B-cell binding by monoclonal antibodies and gp120. *b*, Binding of EBV-transformed MHC class II antigen loss mutant B-cell lines to CD4-transfected COS-1 cells.

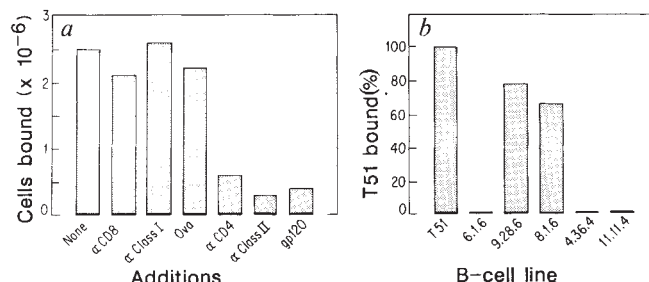
METHODS. For calculation of the number of B cells bound, 10^8 B cells were labelled for 2 h at 37 °C in 0.5 to 1.0 mCi of ⁵¹Cr, washed four times and then used for binding, as described in the legend to Fig. 1. An aliquot of labelled cells was used to determine the specific activity which ranged from 0.01 to 0.1 c.p.m. per B cell. After washing the unbound B cells from individual wells, 1 ml PBS containing 1% Triton was added to each well. Cells were incubated 15 min at 37 °C, lysates spun to remove cell debris and 100 μ l supernatant counted to determine the number of B cells bound to transfected COS-1 cells. Each B-cell line was assayed for binding in duplicate wells of CD4-transfected COS-1 cells in two independent experiments. Counts bound to CDM8-transfected well were subtracted from counts bound to CD4-transfected wells. For T51, c.p.m. bound to CD4-transfected COS-1 cells were always 2-3 times the c.p.m. bound to CDM8-transfected COS-1 cells. For antibody-inhibition studies, monoclonal antibodies were added at a 1:100 dilution to B cells and transfected COS-1 cells, the cells incubated 30-60 min at 37 °C, then mixed and assayed as described in the legend to Fig. 1. Monoclonal antibodies used were: anti-CD4, OKT4; anti-CD8, 7Pt3F9; anti-class I, W6/32; anti-class II, 9/49. Recombinant gp120 from the H3DCG isolate of HIV-1 (Genentech) and ovalbumin (Sigma) were added to a final



OKT4 and 19Thy5D7, respectively, and purified MT321 at 20 μ g ml⁻¹) and analysed after staining with a 1:40 dilution of FITC-conjugated goat anti-mouse immunoglobulin (Mely) on an Epics V cell sorter (Coulter). Propidium iodide was included to gate out dead cells. The anti-CD2 monoclonal antibody 3T4-8B5 was used as a negative control. For gp120 binding, cells were scraped from the well, incubated with 1 μ g gp120 in 0.1 ml for 30 min on ice, stained with 200 ng monoclonal anti-gp120 antibody (Dupont) and FITC-conjugated second antibody and analysed by FACS as above. As a negative control for anti-gp120 antibody staining, no gp120 was added.

for the M11 mutant. M1.1, M1.2 and M5 also gave a reduced B-cell binding phenotype. The mutants M1B, M3 and M8 were indistinguishable from vector-only transfectants (Fig. 1*a*, right panel). The other mutations were without effect.

In contrast to the large number of mutations that affect class II MHC binding, only M3 abrogated gp120 binding (Fig. 1*b* and Table 1). Note that mutants M6, M9 and M14 could not be evaluated as their alterations grossly affect the structure of the external CD4 domains, so that reactivity with all our anti-CD4 monoclonal antibodies is reduced or, in the case of M14,



concentration of 20 μ g ml⁻¹ to the B cells and transfected COS-1 cells, the cells incubated 30-60 min at 37 °C, mixed and assayed. MHC-surface expression of the class II antigen-loss mutant EBV-transformed B cell lines was confirmed by FACS using the following monoclonal antibodies; 9/49 binds to a common determinant on DR, DP and DQ (ref. 34), L243 binds DR (ref. 35), and B7/21 (gift from I. Trowbridge) binds DP, Mutant B-cell haplotypes: T51 (parental, DR1, 3; DQ1, 2; DPX, 4) (ref. 36); 6.1.6 (DR-, -; DQ-, -; DP-, -) (ref. 37); 9.28.6 (DR1-, -; DQ1-, -; DPX-, -) (ref. 36); 8.1.6 (DR3-, -; DQ2-, -; DP4, X) (ref. 36); 4.36.4 (DR-, -; DQ1-, -; DPX-, -) (ref. 38); 11.11.4 (DR-, -; DQ1-, -; DP-, -) (ref. 38). In *b*, specific binding was 8,842 c.p.m.

eliminated. All the other 14 mutants, when transfected into COS-1 cells, yielded a copy number of variant CD4 molecules equivalent to the copy number of wild-type CD4, as assessed by quantitative FACS analysis with anti-CD4 monoclonal antibodies OKT4, 19Thy5D7 and/or MT321.

Using the immunoglobulin V κ Bence-Jones homodimer, REI

TABLE 1 Analysis of CD4 mutants for expression, gp120 binding and MHC class II binding

Mutant	Amino-acid change*	Antibody reactivity	gp120 binding	Class II MHC binding
M1.1	17, T to E 18, A to S	+	+	+/-
M1.2	23, S to I 24, I to T 25, Q to V	+	+	+/-
M1B	27, H to T 30, N to F 32, N to D 34, I to R	+	+	-
M2	40, Q to H	+	+	+
M3	48, P to G 50, K to P 51, L to S	+	-	-
M4	64, Q to K	+	+	+
M5	72, K to N 73, N to K	+	+	+/-
M6	80, D to Q	↓	NE	NE
M7	88, D to N 89, Q to R 94, Q to E	+	+	+
M8	99, G to K 104, S to P 107, H to S	+	+	-
M9	121, P to S 122, P to K 123, G to V	↓	NE	NE
M10	127, S to L 128, V to T 129, Q to E	+	+	+
M11	132, S to H 133, P to K 137, N to V	+	+	+/-
M12	143, T to V	+	+	+
M13	150, E to R	+	+	+
M14	155, G to D 156, T to F 158, T to N	-	NE	NE
M15	162, L to T 163, Q to L 164, N to D	+	+	+

The monoclonal antibody reactivity and gp120 binding were determined by FACS analysis as described in the legend to Fig. 1. Symbols: +, antibody reactivity equivalent to that obtained with wild-type CD4-transfected COS-1 cells; ↓, levels of expression ~2 logs lower; -, no detectable expression. Results are shown for monoclonal antibody OKT4, except in the case of M3 and M8, for which monoclonal antibody MT321 was used. Class II MHC binding was quantitated as described in the legend to Fig. 1. +, Binding of T51 B cells similar to that depicted in Fig. 1a (left panel) and ≥90% of that obtained with wild-type CD4, as depicted by the quantitative method given in Fig. 2 legend; +/-, binding of T51 B cells similar to that depicted in Fig. 1a (middle panel) and 10–20% of that obtained with wild-type CD4; -, no detectable binding, as depicted in Fig. 1a (right panel) and <10% of that obtained with wild-type CD4. Note that no mutants gave class II MHC binding between 20% and 90% of the wild-type CD4. NE, not evaluated because of low expression. * Position, nature of substitution (one-letter code).

(ref. 21), as a model for the CD4 N-terminal domain, the positions of mutants in domain I of CD4 that affect gp120 binding could be assigned to the CDR2 segment¹³. But our analysis of class II MHC binding indicates that a single localized segment is not involved. Rather, residues spread through a substantial portion of the molecule are implicated in CD4-class II interactions. These include amino acids¹³ in the CDR1- and 2-like regions and a segment distal to CDR3 modelled using a published alignment of CD4 and V κ proteins¹⁰. Although some or all of these residues may be contact sites for MHC class II, we cannot exclude the possibility that alteration of residues in these regions affects class II MHC binding to CD4 elsewhere in the molecule. Also, the involvement of regions outside the predicted CDR-like loops such as M5, which is probably located at the opposite pole of the domain from the CDRs^{13,22}, suggests that interaction of CD4 with class II MHC may not be exclusively immunoglobulin-like with respect to binding of ligand by CDR loops. Furthermore, CD4 monomers may interact with one another to create a functional binding site.

Because the CD4 amino-acid residues affecting class II MHC binding are scattered throughout domain I and include at least a stretch in domain II, we speculate that the interactive surface interface between CD4 and class II MHC is much broader than the putative gp120-binding site in CD4. Binding of CD4 with gp120 was only abrogated by the M3 mutation. If both domains I and II are critical for class II MHC binding, then mutations affecting the interaction of these domains with each other might also abolish binding, which could be one effect of mutations outside the regions equivalent to the CDRs. Evidence for the proximity of domains I and II is provided by the fact that antibodies which bind to separate epitopes on these domains cross-block one another's binding to CD4 (ref. 11). It is also possible that there are several discontinuous binding regions for class II MHC.

As M3 abrogated binding of both class II MHC and gp120, we examined the effect of gp120 itself on class II MHC interaction with CD4. Preincubation of B cells with a concentration of gp120 that saturates CD4 binding sites (20 $\mu\text{g ml}^{-1}$) inhibits CD4-class II MHC binding (Fig. 2a). This inhibition of B cell binding to CD4-expressing COS-1 cells is not a result of down-

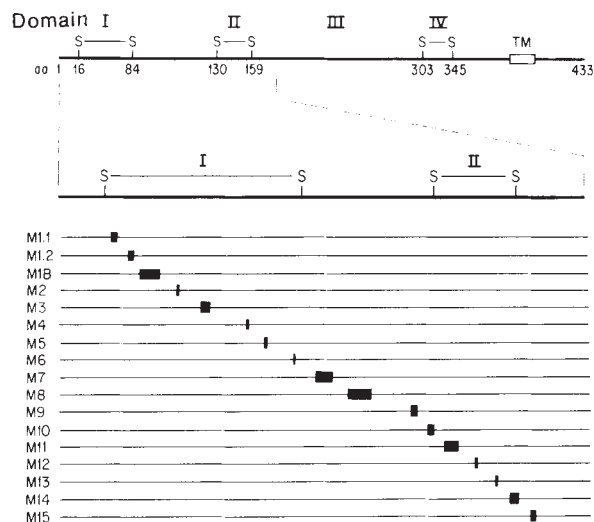


FIG. 3 Positions of mutations in CD4. A schematic diagram of the CD4 protein showing the four immunoglobulin like domains, three disulphide bonds, transmembrane region and cytoplasmic tail. Numbering of amino acids is according to ref. 23. The expanded portion shows the positions of substitutions in 17 different mutants. Mutations were created using the thionucleotide method of oligonucleotide site-directed mutagenesis as described¹³, except that the template for mutagenesis was the full-length CD4 cDNA in M13 phage. The structure of all mutants was confirmed by DNA sequence analysis of the CDM8-CD4 mutant construct using the double stranded DNA as a template.

modulation of CD4 on the transfected cells as their reactivity with OKT4 is unchanged (data not shown). Unlike gp120, soluble CD4(T₄_{ex}) (ref. 23), even at a concentration of 1 mg ml⁻¹, has no effect on binding of B cells, leaving unaffected conjugates such as those seen in Fig. 1a (left panel). This is consistent with the affinity of the monomeric CD4-MHC interaction being much lower (by about ≥ 4 orders of magnitude) than that of CD4 for gp120 (ref. 24), as well as with the lack of inhibition by soluble CD4 of class II-restricted T-cell responses²³. Given the low affinity of monomeric CD4 for class II MHC, the greatly up-regulated expression of CD4 copy number on activated T-lymphocyte clones²⁵ probably enhances interactions of CD4⁺ cells with Ia-expressing cells through an increase in multipoint attachment.

The inhibition of the CD4-class II interaction by gp120 may be significant in T-cell function in HIV-infected individuals in whom gp120 is shed systemically²⁶. CD4⁺ T cells can be depleted by direct lytic infection²⁷, viral cytopathic effects such as syncytia formation^{28,29}, or by a gp120-specific cytolytic mechanism³⁰, but our results have direct implications for the functional activity of the residual normal CD4⁺ T cells. Specifically, gp120 inhibition of CD4 interaction with class II MHC on antigen-presenting cells could inhibit helper T-cell responses which are virtually all class II-restricted, thereby paralysing the normal residual CD4⁺ T lymphocytes. T-cell responses are inhibited by gp120 *in vitro*^{31,32} and our findings provide direct evidence for inhibition of MHC class II-CD4 binding by soluble gp120, which is probably the basis for the observed immunosuppression. Furthermore, it has recently been shown that CD4⁺CD8⁺ precursor T lymphocytes develop into CD4⁺CD8⁻ T lymphocytes as a consequence of interaction between CD4 and class II MHC during thymic development¹, so disruption of this interaction by gp120 might block subsequent production of CD4⁺ T lymphocytes. □

The CD2 antigen associates with the T-cell antigen receptor CD3 antigen complex on the surface of human T lymphocytes

Marion H. Brown*, Doreen A. Cantrell†, Göran Brattsand‡, Michael J. Crumpton* & Martin Gullberg‡

* Cell Surface Biochemistry, † Lymphocyte Activation Laboratories, Imperial Cancer Research Fund, Lincoln's Inns Fields, London WC2A 3PX, UK

‡ Unit for Applied Cell & Molecular Biology, University of Umeå, S-90187 Umeå, Sweden

T LYMPHOCYTES can be activated by various stimuli directed either against the T-cell antigen receptor-CD3 antigen complex (Ti-CD3) or the CD2 molecule; see ref. 1 for a review. Activation signals generated by antigen binding to the antigen-specific α/β heterodimer (Ti) are thought to be transduced via the invariant CD3 γ , ϵ and δ chains, and the associated ζ and η subunits^{2,3}. The physiological role of the interaction of CD2 with its homologous cell-surface associated ligand LFA-3^{4,5} remains to be fully elucidated. It has been suggested that CD2 regulates an antigen-independent pathway of activation⁶ or that signals delivered via CD2 are an integral part of the antigen-specific pathway⁷⁻¹⁰. Several recent studies have indicated a requirement for the Ti-CD3 complex in CD2 signalling. Thus, mutant T-cell lines expressing CD2, but not Ti-CD3, on the cell surface cannot be activated via the CD2 molecules^{9,10}. Functional interaction between the Ti-CD3 complex and the CD2 antigen suggests that these T-lymphocyte cell-surface structures are physically associated. Here we use a digitonin-based solubilization procedure to explore this possibility and show that 40% of the cell-surface CD2 molecules can be specifically co-precipitated in association with the Ti-CD3 complex.

Immunoprecipitates of the Ti-CD3 complex were prepared from normal T lymphoblasts labelled at the surface by lactoperoxidase-catalysed iodination and lysed in digitonin to preserve noncovalent associations. SDS-PAGE analysis under non-reducing conditions of an immunoprecipitate, which had been prepared using UCHTI, a monoclonal antibody against CD3, revealed the characteristic γ -, δ - and ϵ -chains of CD3, plus a band migrating with a relative molecular mass M_r of 90,000 (90K) comprising the disulphide-linked α - and β - chains of Ti (Fig. 1, lane 1). In addition, a diffuse band at 50-55K was clearly visible which showed remarkable similarity in mobility and appearance to the authentic CD2 immunoprecipitate shown in lane 2. Under reducing conditions, the α -chain of Ti has an M_r of 50-55K, making it difficult to identify any other molecule of similar mobility in the CD3 immunoprecipitate (lane 4). The appearance of CD2 in the CD3 immunoprecipitate represents genuine co-precipitation and was not due to cross-reactivity of UCHTI with CD2. This was demonstrated by preparing immunoprecipitates using UCHTI from digitonin lysates of insect Sf9 cells expressing high levels of CD2 after infection with a recombinant baculovirus. Under these circumstances, no specific precipitation of CD2 by UCHTI was detected (data not shown).

To provide evidence that the 50-55K band which coprecipitated with the Ti-CD3 complex was the CD2 antigen, we performed a depletion experiment. A digitonin lysate, depleted of CD2 antigen by pre-clearing with a monoclonal antibody against CD2 (lane 8), was reprecipitated with a monoclonal antibody against CD3. Analysis of this CD3 precipitate under non-reducing conditions (lane 10) revealed no band at 50-55K, suggesting that the band observed in this region in lane 1 did

Received 6 March; accepted 10 May 1989.

- Kruisbeek, A. M. *et al. J. exp. Med.* **161**, 1029-1047 (1985).
- Krensky, A. M., Reiss, C. S., Mier, J. W., Strominger, J. L. & Burakoff, S. J. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2365-2369 (1982).
- Meuer, S. C., Schlossman, S. F. & Reinherz, E. L. *Proc. natn. Acad. Sci. U.S.A.* **79**, 4395-4399 (1982).
- Biddison, W., Rao, P., Talle, M. A., Goldstein, G. & Shaw, S. J. *J. exp. Med.* **156**, 1065-1076 (1982).
- Marrack, P. *et al. J. exp. Med.* **158**, 1077-1091 (1983).
- Klatzmann, D. *et al. Science* **225**, 59-63 (1984).
- Dalgleish, A. G. *et al. Nature* **312**, 763-766 (1984).
- Sattentau, Q., Dalgleish, A., Weiss, R. & Beverley, P. C. L. *Science* **234**, 1120-1123 (1986).
- McDougal, J. S. *et al. Science* **231**, 382-385 (1986).
- Maddon, P. J. *et al. Cell* **47**, 333-348 (1986).
- Peterson, A. & Seed, B. *Cell* **54**, 65-72 (1988).
- Landau, N. R., Warton, M. & Littman, D. R. *Nature* **334**, 159-162 (1988).
- Clayton, L. K. *et al. Nature* **335**, 363-366 (1988).
- Doyle, C. & Strominger, J. L. *Nature* **330**, 256-259 (1987).
- Eichmann, K., Jonsson, J. I., Falk, I. & Emmrich, F. *Eur. J. Immun.* **17**, 643-650 (1987).
- Rudd, C. E., Trevillyan, J. M., Dasgupta, J. D., Wong, L. L. & Schlossman, S. F. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5190-5194 (1988).
- Veillette, A., Bookman, M. A., Horak, E. M. & Bolen, J. B. *Cell* **55**, 301-308 (1988).
- Mellins, E., Woelfel, M. & Pious, D. *Hum. Immun.* **18**, 211-223 (1987).
- Gay, D., Buus, S., Pasternak, J., Kappler, J. & Marrack, P. *Proc. natn. Acad. Sci. U.S.A.* **85**, 5629-5633 (1988).
- Maddon, P. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 9155-9159 (1987).
- Epp, O., Lattman, E. E., Schiffer, M., Huber, R. & Palm, W. *Biochemistry* **14**, 4943-4952 (1975).
- Richardson, N. E. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 6102-6106 (1988).
- Hussey, R. E. *et al. Nature* **331**, 78-81 (1988).
- Lasky, L. *et al. Cell* **50**, 975-985 (1987).
- Meuer, S. C. *et al. Nature* **303**, 808-810 (1983).
- Schneider, I., Kaaden, O., Copeland, J. D., Orosian, S. & Hunsman, G. *J. gen. Virol.* **67**, 2533-2539 (1986).
- Samasundaran, M. & Robinson, H. L. *J. Virol.* **61**, 3114-3119 (1987).
- Sodroski, J., Goh, W. C., Rosen, C., Campbell, K. & Haseltine, W. A. *Nature* **322**, 470-474 (1986).
- Lifson, J. *et al. Nature* **323**, 725-728 (1986).
- Siliciano, R. F. *et al. Cell* **54**, 561-575 (1988).
- Diamond, D. C. *et al. J. Immun.* **141**, 3715-3717 (1988).
- Weinhold, K. J. *et al. J. Immun.* (in press).
- Seed, B. & Aruffo, A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 3365-3369 (1987).
- Pesando, J. M. & Graf, L. *J. Immun.* **136**, 4311-4318 (1986).
- Lampson, L. A. & Levy, R. *J. Immun.* **125**, 293-299 (1980).
- Levine, F. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 3741-3745 (1985).
- Levine, F., Erlich, H., Mach, B. & Pious, D. *J. Immun.* **134**, 637-640 (1985).
- Pious, D., Dickson, L., Levine, F., Cotner, T. & Johnson, R. J. *J. exp. Med.* **162**, 1193-1207 (1985).

ACKNOWLEDGEMENTS. We thank Professor Stephen C. Harrison for advice and assistance with computer modelling. Drs George Tarr, Shigeo Koyasu and Frederick Cohen for critical reading of this manuscript and Dr Gordon Freeman for advice on COS-1 cell transfections. This work was supported in part by the NIH.

indeed represent CD2. CD2, was, however, precipitated (lane 9) from a lysate that had been depleted of the Ti-CD3 complex (lane 7), implying that if the original 50-55K band was CD2, then not all of it was associated with Ti-CD3. Densitometric analysis of the experiment shown in Fig. 1 (lanes 1 and 9) showed that ~40% of the CD2 present in the lysate was co-precipitated with the Ti-CD3 complex.

The identity of the cell-surface iodinated 50-55K band immunoprecipitated by a monoclonal antibody against CD3 was further investigated by an immunoblotting experiment. The iodinated proteins on the gel shown in Fig. 1 were electrophoretically transferred to a nitrocellulose filter. A 16-h exposure of the transfer is shown before (Fig. 2a) and after (Fig. 2b) detection of CD2 by using a rabbit antiserum against purified denatured human CD2, the specificity of which has been reported previously^{11,12} and which provides a very sensitive means of detecting CD2 in western blot analysis. Thus, the experiment shown in Fig. 2 unequivocally established the identity of the 50-55K molecule, seen in the unreduced CD3 immunoprecipitate (Fig. 1, lane 1), as CD2. CD2 was also identified with a directly iodinated monoclonal antibody against CD2 called RFTII¹², by blotting CD2 and CD3 immunoprecipitates on transfers containing increased amounts of antigen prepared from an unlabelled lysate (data not shown).

CD2 antigen was also identified by western blot analysis in CD3 immunoprecipitates prepared from digitonin lysates of the T-leukaemia cell line Jurkat (Fig. 3a, lane 1). The physical association of CD2 with the Ti-CD3 complex appears to be specific in that the CD2 antigen was not detected in immunoprecipitates, prepared under the same conditions, of several other cell-surface antigens including the interleukin-2 (anti-Tac) and transferrin receptors (Fig. 3a, lane 3; Fig. 3b,

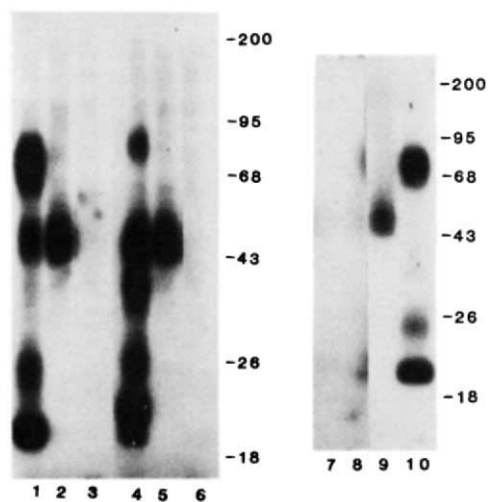


FIG. 1 Co-precipitation of the CD2 antigen with the Ti-CD3 complex. Immunoprecipitates from T lymphoblasts iodinated at the surface and solubilized in digitonin were prepared with UCHT1 (lanes 1 and 4); a monoclonal antibody against CD2, OKT11 (lanes 2 and 5); both rabbit and mouse normal immunoglobulin (lanes 3 and 6) and analysed under non-reducing (lanes 1-3) or reducing (lanes 4-6) conditions. Supernatants from precipitates shown in lanes 1 and 2 were reprecipitated with UCHT1 (lane 7) or OKT11 (lane 8), respectively, to show the depletion of antigen and then with the reciprocal antibody OKT11 (lane 9) and UCHT1 (lane 10), and were analysed under non-reducing conditions.

METHODS. T lymphoblasts, prepared as described¹⁴, were labelled at the surface by lactoperoxidase-catalysed iodination (Na^{125}I , IMS30; Amersham) and lysed at 10^7 cells per ml in 10 mM Tris-HCl, pH 7.2, 0.15 M NaCl containing 1% digitonin and protease inhibitors, including 20 mM iodoacetamide²². Immunoprecipitates were prepared²² by incubation for 16 h at 4 °C with antibody directly coupled to Sepharose 4B. Precipitates were washed twice in 10 mM Tris-HCl, pH 7.4, 0.15 M NaCl containing 0.05% digitonin and once in 10 mM Tris-HCl, pH 7.4, before analysis on a 15% SDS polyacrylamide gel, fixation in 50% methanol, 7% acetic acid and autoradiography²². M_r of standards in thousands are shown on the right.

lanes 3 and 4), as well as the class I and LFA-1 antigens (data not shown). In contrast to immunoprecipitation with the CD3 antibody, CD2 antibodies were not efficient in co-precipitating the Ti-CD3 complex. Although evidence for putative Ti-CD3 chains associated with a CD2 immunoprecipitate was obtained in several experiments, the observation was not consistently reproducible for reasons which are unclear.

The demonstration of co-precipitation between CD2 and Ti-CD3 was dependent on the detergent used for cell lysis. No evidence was obtained for CD2 in a CD3 immunoprecipitate prepared from cells solubilized in Triton X100 (Fig. 3c, lane 1), which may explain why the phenomenon has not been reported previously. The CD3 immunoprecipitate prepared from the comparable digitonin lysate (Fig. 3c, lane 2), revealed two extra diffuse bands, one at ~50K which is probably monomeric CD2 and another band at about 100K which disappeared on reduction (Fig. 3b, lane 1). Because the M_r of the latter band is twice that of monomeric CD2, this band could represent CD2 homodimer, direct evidence for which was provided by immunoblotting experiments. Thus, the rabbit anti-CD2 serum identified a 100K band in an unreduced CD2 immunoprecipitate (Fig. 3d). The relative amounts of dimeric (18%) and monomeric (82%) CD2 shown in this blot (Fig. 3d) are more representative of many experiments than that shown in Fig. 3c. Although lysis was carried out in the presence of iodoacetamide, the dimer could represent an artefact of disulphide cross-linkage, as has been previously suggested for class I antigens¹³. However, with an estimated stoichiometry of six molecules (that is, more than one) of CD2 associated with one Ti-CD3 complex, the ability of CD2 to dimerize could be of physiological relevance. This stoichiometry is derived from the observation that T lymphoblasts express ~250,000 molecules of CD2 antigen⁶ and that ~40% of this is associated with the 17,000 molecules of Ti-CD3 on these cells¹⁴.

Non-covalent associations of the Ti-CD3 complex have been reported with a number of other T-cell surface antigens, in particular CD4 and CD8. Evidence in support of such associations has been provided by both co-modulation experiments

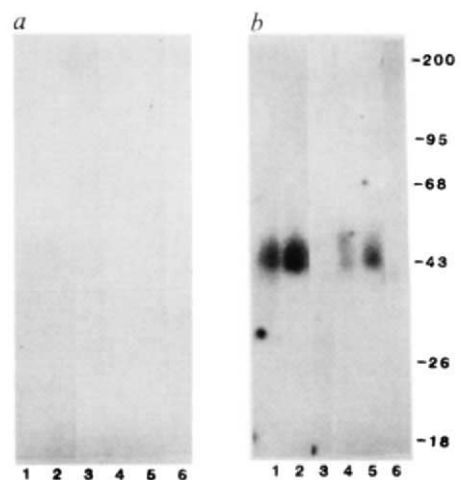
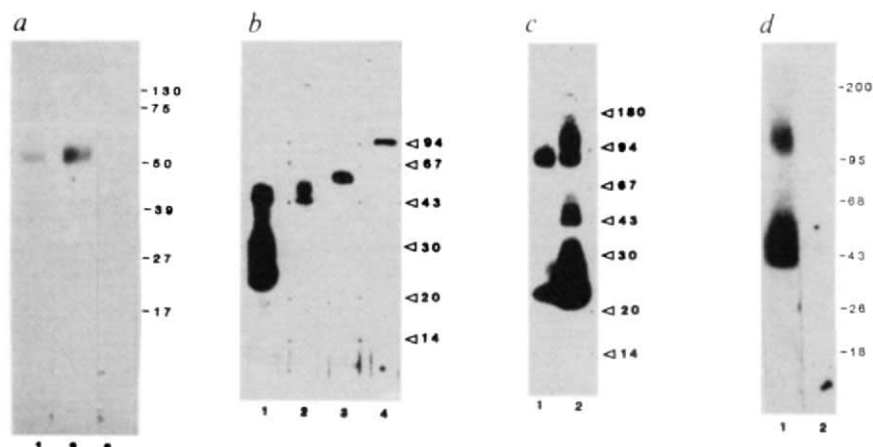


FIG. 2 Immunoblotting of the Ti-CD3 immunoprecipitates with a rabbit anti-CD2 serum. The gel containing lanes 1-6 from Fig. 1 was rehydrated in 25 mM Tris-192 mM glycine, pH 8.3 containing 20% (v/v) methanol and the proteins transferred electrophoretically for 24 h at 50 v to Hybond C extra (Amersham). a, 16 h exposure of transfer. b, 16 h exposure of transfer after detection of CD2 with rabbit antiserum against denatured purified human CD2^{11,12}. CD2 was detected first by incubating the transfer firstly with 5% dried skim milk powder and 2% bovine serum albumin (BSA) in phosphate buffered saline (PBS), pH 7.2 for 2 h at 22 °C; second, with a rabbit anti-CD2 serum (1:100 dilution); and third, with ^{125}I -labelled protein A at 0.2 $\mu\text{Ci ml}^{-1}$ (IM 144; Amersham) each diluted in 0.1% BSA in PBS and incubated with the filter for 16 h at 4 °C. After each step the transfer was washed in PBS containing 0.05% Tween and for autoradiography the moist transfer was wrapped in Saranwrap.

FIG. 3 *a*, Coprecipitation of the CD2 antigen with the Ti-CD3 complex in Jurkat cells. Immunoprecipitates prepared with UCHT1 (lane 1); the monoclonal antibody against CD2 MAR206 (lane 2); or against the transferrin receptor OKT9 (lane 3), were analysed under non-reducing conditions. *b*, Specificity of the association between the CD2 antigen and the Ti-CD3 complex. Immunoprecipitates prepared with UCHT1 (lane 1), OKT11 (lane 2), antibody against the IL-2 receptor Tac (lane 3), or OKT9 (lane 4) were analysed under reducing conditions. *c*, Solubilization conditions. UCHT1 immunoprecipitates prepared from cells lysed at 10^7 cells per ml in 1% Triton X100 (lane 1) or digitonin (lane 2) were analysed under non-reducing conditions. *d*, Immunoblot of dimeric CD2 antigen. An unreduced OKT11 immunoprecipitate from unlabelled T lymphoblasts was immunoblotted with the rabbit anti-CD2 serum^{11,12} (lane 1) or mouse normal serum (lane 2) as described in Fig. 2 legend.



METHODS. Immunoprecipitates were prepared from the T-leukaemia cell-line Jurkat (*a*) or from T lymphoblasts generated by stimulation with UCHT1 or OKT3 (*b, c*), or phytohaemagglutinin (*d*) as described in Fig. 1 legend and

and the co-stimulatory properties of specific monoclonal antibodies in the induction of a mitogenic response¹⁵. Direct evidence, such as co-immunoprecipitation, has still not been reported. Here we show that ~40% of the CD2 antigen expressed at the cell surface can be specifically co-precipitated with the Ti-CD3 complex. These interactions might appear unexpected as it has been shown that modulation of Ti-CD3 does not decrease CD2 expression^{14,16}. The relative strength of the interaction may not, however, be sufficient for co-modulation, as the interaction between CD2 and Ti-CD3 appears less stable than that between Ti and CD3 (Fig. 3c). Association between CD2 and the Ti-CD3 complex may explain the results of experiments in which simultaneous binding of specific ligands to CD2 and Ti-CD3 transmits an activation signal to the cell, under conditions where occupancy of one receptor is insufficient^{15,17}. In cases where it appears that CD2 transmits an activation signal in the absence of Ti-CD3 expression^{18,19}, CD2 could function by associating with an unidentified receptor component.

The mechanism of association between CD2 and Ti-CD3 needs to be explored and the relevant interacting regions of the molecules identified. A region of amino-acid sequence similarity between the membrane proximal extracellular domains of CD2 and CD4 has previously been noted^{11,20}. It may be that both CD4 and CD2 interact at these conserved sites with Ti-CD3. Alternatively, as the CD2 cytoplasmic domain is necessary for signal transduction^{17,21}, this domain could interact with Ti-CD3. It will be interesting to determine whether CD2 mutants lacking a functional cytoplasmic domain associate with Ti-CD3. □

analysed on 11% (*a*), 10–20% (*b, c*) or 15% (*d*) SDS-PAGE. (*a–c*, Standards from Pharmacia or *d*, Bethesda Research Lab.)

20. Williams, A. F., Barclay, A. N., Clark, S. J., Paterson, D. J. & Willis, A. C. *J. exp. Med.* **165**, 368–380 (1987).

21. He, Q., Beyers, A. D., Barclay, A. N. & Williams, A. F. *Cell* **54**, 979–984 (1988).

22. Davies, A. A. & Brown, M. H. in *Lymphocytes: Practical Approach Series*, (ed. Klaus, G.) 229–253 (IRL, Oxford, 1987).

ACKNOWLEDGEMENTS. We thank Susan Lucas for help with tissue culture, Dr Mike Owen for review of the manuscript and Maureen Stilwell for secretarial assistance. Supported in part by the Swedish Medical Research Council.

Regulation of pro-opiomelanocortin biosynthesis and processing by transplantation immunity

S. Zakarian, M. S. Eleazar & W. K. Silvers*

The Wistar Institute, Thirty-sixth Street at Spruce, Philadelphia, Pennsylvania, 19104 USA

* Department of Human Genetics, The University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania, 19104, USA

IT is more than thirty years since Billingham and Medawar showed that adrenocorticotrophic hormone (ACTH) and cortisol can prolong the survival of skin allografts^{1,2}. It has since become clear that glucocorticoid hormones are critically involved in the regulation of immunity^{3,4}. The level of glucocorticoids secreted in response to antigenic challenge corresponds to the magnitude of the immune response and in general reaches immunosuppressive levels^{5–7}. Interestingly, not all immune responses enhance ACTH and glucocorticoid hormone production. In transplantation immunity, the reverse seems to be true: circulating glucocorticoid levels at the time of skin graft rejection are lower than control levels⁸. Because β -endorphin and ACTH originate from the same prohormone, pro-opiomelanocortin (POMC)⁹, and are closely related in their tissue-specific processing and coordinate release^{10–12}, we have investigated the role of pituitary β -endorphin in transplantation immunity. We report here that POMC biosynthesis and processing in the pars intermedia, but not in the anterior pituitary, can be regulated by T cell-specific factors secreted in animals undergoing transplantation immunity.

We grafted Lewis rats¹³ with either isogenic skin or genetically unrelated and major histocompatibility complex-incompatible DA rat skin on one side of the thorax and, two weeks later, with an identical graft on the other side of the thorax. The first allografts were rejected in 8–9 days, whereas the second rejection was accelerated (5–6 days), a characteristic of the secondary immune response. Isogenic grafts were accepted indefinitely.

Received 15 February; accepted 8 May 1989.

- Möller, G. (ed) *Immun. Rev.* **95** (Munksgaard, Copenhagen, 1987).
- Clevers, H., Alarcon, B., Wiseman, T. & Terhorst, C. *A. Rev. Immun.* **6**, 629–662 (1988).
- Baniyash, M., Garcia-Morales, P., Bonifacio, J. S., Samelson, L. E. & Klausner, R. D. *J. biol. Chem.* **263**, 9874–9878 (1988).
- Hünig, T. *J. exp. Med.* **162**, 890–901 (1985).
- Selvaraj, P. *et al. Nature* **326**, 400–403 (1987).
- Meuer, S. C. *et al. Cell* **36**, 897–906 (1984).
- Yang, S. Y., Chouaib, S. & Dupont, B. *J. Immun.* **137**, 1097–1100 (1986).
- Breitmeyer, J. B., Daley, J. F., Levine, H. B. & Schlossman, S. F. *J. Immun.* **139**, 2899–2905 (1987).
- Alcover, A. *et al. EMBO J.* **7**, 1973–1977 (1988).
- Bockenstedt, L. K. *et al. J. Immun.* **143**, 1904–1911 (1988).
- Sewell, W. A., Brown, M. H., Dunne, J., Owen, M. J. & Crumpton, M. J. *Proc. natn. Acad. Sci. U.S.A.* **83**, 8718–8722 (1986).
- Brown, M. H., Sewell, W. A., Monostori, E. & Crumpton, M. J. in *Leucocyte Typing III* (ed. McMichael, A. J.) 110–112 (Oxford Univ. Press 1987).
- Snary, D., Goodfellow, P., Bodmer, W. F. & Crumpton, M. J. *Nature* **258**, 240–242 (1975).
- Cantrell, D. A., Davies, A. A. & Crumpton, M. J. *Proc. natn. Acad. Sci. U.S.A.* **82**, 8158–8162 (1985).
- Anderson, P., Morimoto, C., Breitmeyer, J. B. & Schlossman, S. F. *Immun. Today* **9**, 199–203 (1988).
- Cantrell, D. A., Verbi, W., Davies, A., Parker, P. & Crumpton, M. J. *Eur. J. Immun.* **18**, 1391–1396 (1988).
- Bierer, B. E., Peterson, A., Gorga, J. C., Herrmann, S. H. & Burakoff, S. J. *J. exp. Med.* **168**, 1145–1156 (1988).
- Pantaleo, G. *et al. J. exp. Med.* **168**, 13–24 (1988).
- Siliciano, R. F., Pratt, J. C., Schmidt, R. E., Ritz, J. & Reinherz, E. L. *Nature* **317**, 428–430 (1985).

TABLE 1 Peptides with β -endorphin/ACTH immunoreactivity in regions of normal and immunized Lewis rat pituitary

Tissue/column	β -Endorphin/ACTH-related peptides	β -Endorphin/ACTH immunoreactivity (pmol per rat)		
		Normal Lewis	Lewis (isograft)	Lewis anti-DA (allograft)
Anterior pituitary/ Gel filtration G-75	β -lipotropin	73	22 (−70%)*	71 (−3%) [†]
	β -endorphin size peptides	63	26 (−58%)	75 (18%)
	ACTH	102	34 (−67%)	106 (4%)
Pars intermedia and posterior pituitary/ gel filtration G-75	β -endorphin size peptides	1,591	1,521 (−4%)	6,598 (315%)
	ACTH (CLIP)	1,090	1,027 (−6%)	4,497 (313%)
Pars intermedia and posterior pituitary/ ion exchange SP C-25	Ac β -endorphin 1–26	505	688 (36%)	275 (−45%)
	Ac β -endorphin 1–27	525	334 (−36%)	633 (21%)
	β -endorphin 1–26	45	92 (102%)	105 (131%)
	β -endorphin 1–27	90	97 (8%)	58 (−35%)
	Ac β -endorphin 1–31	199	170 (−14%)	2,020 (916%)
	β -endorphin 1–31	208	114 (−45%)	3,402 (1,537%)

* Per cent change in immunoreactive β -endorphin/ACTH-related peptides in isografted rats as compared with normal Lewis rat peptides: $(22 - 73)/73 = -70\%$.

[†] Per cent change in immunoreactive β -endorphin/ACTH-related peptides in allografted rats as compared with normal Lewis rat peptides: $(71 - 73)/73 = -3\%$.

Nine days after the second graft challenge, we dissected the pituitaries from isografted and allografted Lewis rats, and those from ungrafted littermates, into the anterior pituitary and the pars intermedia plus posterior pituitary. We extracted the tissues and resolved the ACTH and β -endorphin-related peptides by gel filtration and ion-exchange chromatography in the presence of trace amounts of iodinated marker peptides^{14,15}. Endogenous peptides were resolved by column chromatography and measured by radioimmunoassay using well-characterized antisera to ACTH and β -endorphin (Fig. 1).

Gel filtration and immunoassay of the peptides extracted from the regions of the pituitary of normal, isografted and allografted Lewis rats all gave the normal patterns of POMC-processing characteristic of each region of the pituitary: ACTH, β -

lipotropic hormone (LPH) and β -endorphin immunoreactivity in the anterior pituitary, and melanocyte-stimulating hormone (α -MSH), corticotropin-like intermediate peptide (CLIP) and β -endorphin immunoreactivity in the pars intermedia (Fig. 1). The amount of immunoreactive peptides in the anterior pituitary of normal and allografted rats was comparable, but in the isografted rats, immunoreactive LPH, β -endorphin and ACTH were substantially lower. This was probably a consequence of a regulatory mechanism of wound-healing involving ACTH and glucocorticoid actions. Immunoreactive peptides in the pars intermedia of the control and transplanted rats showed there was a threefold increase in CLIP and β -endorphin immunoreactivity in the allografted animals only (Table 1). The results indicate a specific increase in POMC biosynthesis which

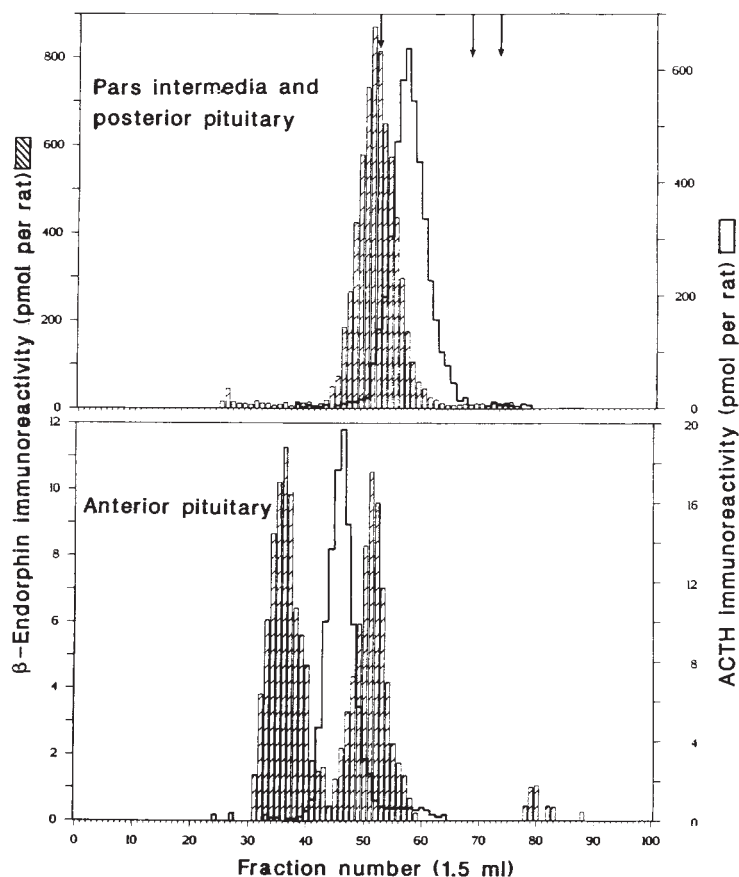


FIG. 1 Resolution by gel filtration of β -endorphin and ACTH-related peptides in the pars intermedia plus posterior pituitary (top) and in the anterior pituitary (bottom) of allografted Lewis rats. In the pars intermedia, the hatched peak represents immunoreactive β -endorphin-sized peptides, and the blank peak represents immunoreactive CLIP. In the anterior pituitary, the hatched peaks from left to right represent LPH and β -endorphin immunoreactivity, and the blank peak represents ACTH immunoreactivity.

METHODS. The regions of pituitary from normal and immunized rats were extracted in acetone/HCl/water (40:1:5), and the soluble products were fractionated by gel filtration on a Sephadex G-75 column (70 $\times$ 1.5 cm) in 50% acetic acid^{14,15}. The peak elution position of the group of β -endorphin-related peptides was identified by adding ¹²⁵I-labelled camel marker peptides (Peninsula Laboratories) to each homogenate. The iodinated camel marker peptides are β -endorphin 1–31, β -endorphin 1–27, β -endorphin 1–26, and their α ,N-acetyl forms. Aliquots from the pars intermedia plus posterior pituitary (5 μ l) and from the anterior pituitary (20 μ l) were dried and radioimmunoassayed with a well-characterized β -endorphin antiserum¹⁹. A duplicate set of samples from the same column were assayed for ACTH immunoreactivity using an antibody with specificity for the C-terminal sequence of ACTH. The arrows from left to right indicate peak elution positions of ¹²⁵I-labelled β -endorphin, ¹²⁵I-labelled α -MSH and ¹²⁵I-labelled (des acetyl)-MSH respectively. Each experiment used a pool of two Lewis rats and was repeated at least twice.

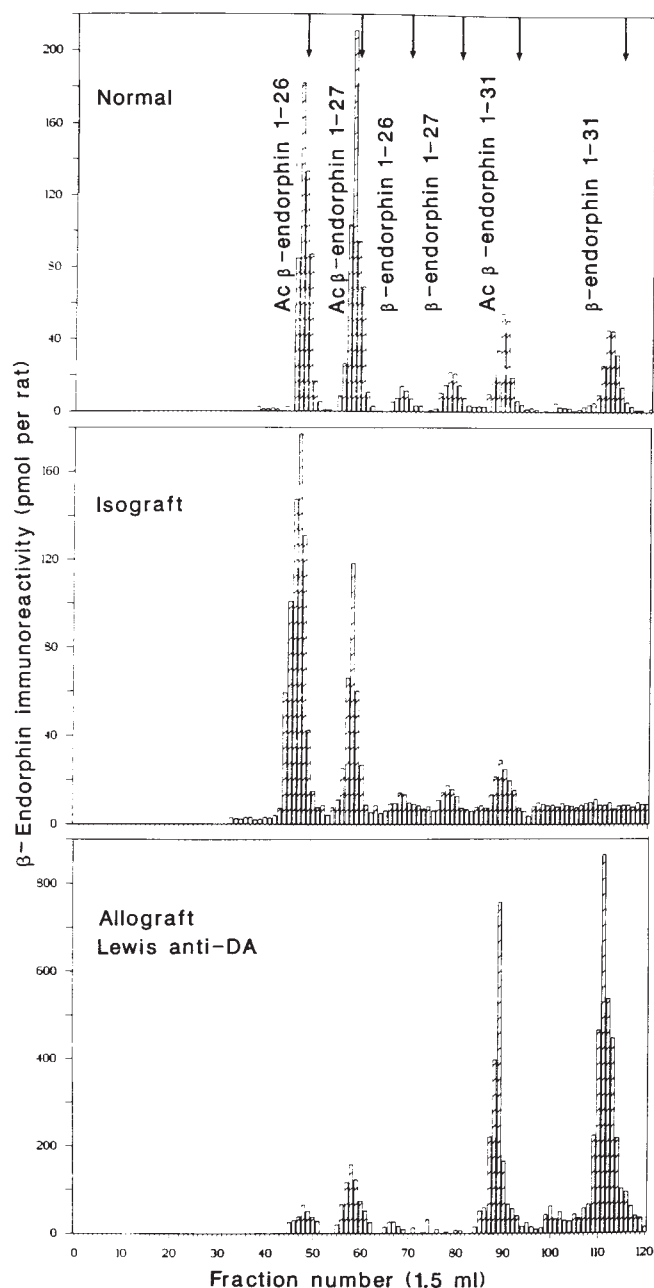


FIG. 2 Resolution by ion-exchange chromatography of β -endorphin-related peptides in the pars intermedia plus posterior pituitary of normal Lewis rats (top), Lewis rats grafted with isogeneic skin (isograft) (middle), and Lewis rats grafted with allogeneic DA skin (allograft) (bottom). Each experiment represents a pool of two Lewis rats and was repeated at least twice. The β -endorphin fraction obtained by gel filtration and containing the six camel β -endorphin marker peptides was chromatographed on a column (50×0.7 cm) of SP-Sephadex C-25 (pyridinium form)¹⁵. Elution was by a linear gradient (100 ml mixer volume) of 0–1 M pyridine in 50% acetic acid. The peak elution positions of the six radiolabelled camel reference peptides, acetyl β -endorphin 1–26, acetyl β -endorphin 1–27, β -endorphin 1–26, β -endorphin 1–27, acetyl β -endorphin 1–31, and β -endorphin 1–31 are indicated from left to right by arrows; recovery of radioactive marker peptides was >90%. Note the increased amount of immunoreactive peptide biosynthesized in the allografted Lewis rats and the dramatic change in the processing pattern of β -endorphin. Also, in the isografted control Lewis rats, the acetylation and processing of β -endorphin proceeds almost to completion, giving rise to acetyl β -endorphin 1–26 as the prominent immunoreactive peptide. Thus in the absence of transplantation immunity, processing of β -endorphin goes to completion, whereas in active transplantation immunity β -endorphin processing is inhibited and the intact β -endorphin 1–31 is stored and/or secreted.

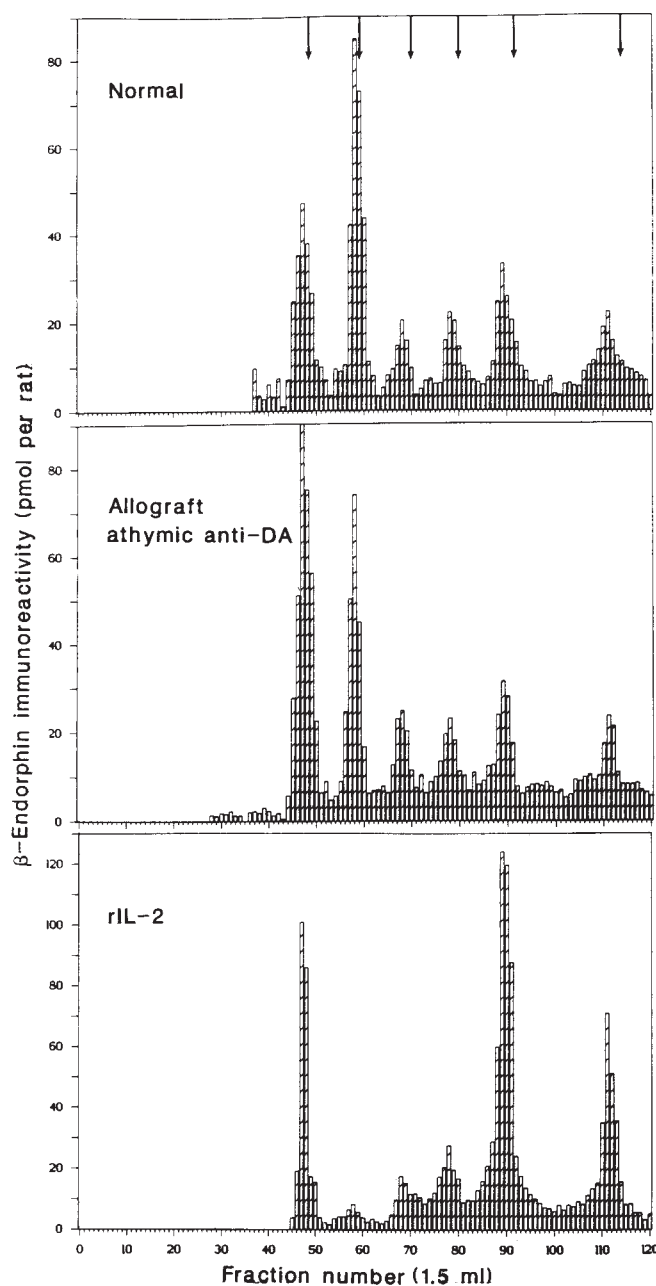


FIG. 3 Resolution by ion-exchange chromatography of β -endorphin-related peptides in the pars intermedia plus posterior pituitary of untreated athymic rats (top), athymic rats grafted with allogeneic DA skin (middle) and athymic rats one hour after an intraperitoneal injection of rIL-2 (bottom). Skin grafting and chromatography are described in Figs 1 and 2. Note the pattern of processing of β -endorphin-related peptides of the allografted rats is similar to that of the normal Lewis rat controls (Fig. 2) and the untreated athymic rat controls. But it is more like the pattern of the Lewis isografted rats (Fig. 2), in which the acetylation and processing of β -endorphin proceeds nearly to completion, giving rise to acetyl β -endorphin 1–26 and not to β -endorphin 1–31 as the main immunoreactive peptide. Also, in rIL-2-injected athymic rats (bottom), the β -endorphin processing in the pars intermedia is highly modulated, revealing a processing pattern similar to the pars intermedia of allografted Lewis rats, in which the intact β -endorphin 1–31 in the acetyl or amino form represents the chief immunoreactive peptide (Fig. 2, bottom). Normal Lewis rats injected with rIL-2 and control rats injected with PBS alone (1 hour after treatment) showed normal processing patterns, similar to the untreated Lewis rat pars intermedia (Fig. 2, top).

corresponds with the height of the secondary immune response in the pars intermedia, but not in the anterior pituitary.

When we chromatographed the immunoreactive β -endorphin gel-filtration peak from normal and isografted rat pars intermedia using ion exchange on SP-Sephadex C-25 (Fig. 2), we saw a similar pattern of β -endorphin-processing in the two groups. The prominent peptides in both cases were the inactive forms of β -endorphin, namely acetyl β -endorphin 1-26 and acetyl β -endorphin 1-27. The remaining four β -endorphin-related peptides, identified from the elution positions of radioiodinated marker peptides, were minor components. The striking result in this experiment is to be seen in the β -endorphin-processing pattern observed in the pars intermedia of the rats after rejection of the second allograft. The processing pattern shifts from the inactive, processed and acetylated β -endorphin 1-26 and 1-27 to acetyl β -endorphin 1-31 and the biologically active β -endorphin 1-31, which becomes the main immunoreactive peptide identified in the pars intermedia (a 15-fold increase over controls: Table 1).

As transplantation immunity is a T cell response, we have considered the possibility that the observed β -endorphin modulation in the pars intermedia is brought about by circulating T cell factors secreted during graft rejection. We therefore grafted NIH congenitally athymic rats with DA skin (allografts) and, as expected, these grafts healed in permanently, demonstrating the lack of an immune response. We also injected athymic rats and normal Lewis rats intraperitoneally with $0.3 \mu\text{g ml}^{-1}$ of recombinant human interleukin-2 (rIL-2) (Collaborative Research Inc.) in phosphate-buffered saline (PBS) and normal Lewis rats with 1 ml of PBS alone. Half an hour after, and 1, 2, 4, 8 and 24 hours after injection, the rats were killed and their pituitaries removed and stored at -70°C .

Identification of the ACTH- and β -endorphin-related peptides from the pituitary regions of both untreated control and transplanted athymic rats indicated that the processing pattern was similar to the control and isografted Lewis rats. Once more, there was a general decrease in the anterior pituitary in the level of immunoreactive peptides, whereas in the pars intermedia the total β -endorphin and ACTH (CLIP) immunoreactivity was normal. But the processing pattern revealed by ion-exchange chromatography showed that the chief immunoreactive peptide stored was the inactive acetyl β -endorphin 1-26, and not the biologically active β -endorphin 1-31 (Fig. 3, top and middle panels). In contrast, the processing pattern of the immunoreactive β -endorphin from the pars intermedia of athymic rats injected with rIL-2 mimics the β -endorphin modulation in allografted rats (Figs 2 and 3, bottom panel). This modulation gives rise to β -endorphin 1-31 and is dependent on the availability of rIL-2 in the circulation, because 4 hours after injecting athymic rats with rIL-2, the β -endorphin-processing pattern reverts to expressing the truncated acetyl β -endorphins 1-26 and 1-27. Furthermore, normal Lewis rats injected with rIL-2 do not show a strong β -endorphin modulation in the pars intermedia, indicating that exogenous rIL-2 is efficiently removed from the circulation by IL-2 receptors on T cells. At the height of the immune response to transplantation, T cells secrete IL-2 in larger amounts than can be absorbed by specific IL-2 receptors, and consequently we observe β -endorphin modulation in allografted rats.

It is important to note that as neither intact POMC nor LPH is stored in the pars intermedia, then the β -endorphin 1-31 immunoreactivity observed in the amino or acetyl forms after allograft rejection or rIL-2 injection can be attributed to newly synthesized products, C-terminal proteolysis and N-terminal acetylation of β -endorphin both being irreversible post-translational modifications¹⁶. Control Lewis rats injected intraperitoneally with PBS alone revealed the normal processing patterns characteristic for both regions of the pituitary, so the mechanisms regulating stress-induced β -endorphin modulation in the pars intermedia¹⁷⁻¹⁸ do not seem to be activated during

transplantation immunity of allografted rats or rIL-2-injected rats. We have demonstrated here that β -endorphin modulation in the pars intermedia of transplanted rats is a T cell and an IL-2 dependent phenomenon. Finally, whereas anterior pituitary ACTH may promote the survival of allografts, β -endorphin from the pars intermedia may be important in the regulation of transplantation immunity and in T cell-dependent immune responses generally. □

Received 1 February; accepted 14 April 1989.

1. Billingham, R. E. & Krohn, P. L. *Br. Med. J.* **1**, 1157-1163 (1951).
2. Medawar, P. B. & Sparrow, E. M. *J. Endocr.* **14**, 240-256 (1956).
3. Besedovsky, H. O., del Rey, A. & Sorkin, E. in *Pituitary Function and Immunity* (ed. Berczi, E.) 241-250 (CRC Boca Raton, 1986).
4. Kelso, A. & Munck, A. *J. Immun.* **133**, 784-791 (1984).
5. Berkenbosch, F., van Oers, J., del Rey, A., Tilders, F. & Besedovsky, H. *Science* **238**, 524-526 (1987).
6. Sapolsky, R., Rivier, C., Yamamoto, G., Plotsky, P. & Vale, W. *Science* **238**, 522-524 (1987).
7. Bernton, E. W., Beach, J. E., Holaday, J. W., Smallridge, R. C. & Fein, H. G. *Science* **238**, 519-521 (1987).
8. Besedovsky, H. O., Sorkin, E. & Keller, M. *J. Endocr.* **76**, 175-176 (1978).
9. Nakanishi, S. *et al. Nature* **278**, 423-427 (1979).
10. Zakarian, S. & Smyth, D. G. *Biochem. J.* **202**, 561-571 (1982).
11. Mains, R. E., Eipper, B. A., Glembofski, C. C. & Dores, R. M. *Trends Neurosci.* **6**, 229-235 (1983).
12. Dores, R. M., Akil, H. & Watson, S. J. *Peptides* **5**, 9-17 (1984).
13. *Transplantation of Tissues and Cells* (eds Billingham, R. E. & Silvers, W. K.) (The Wistar Institute Press, Philadelphia, 1961).
14. Smyth, D. G. & Zakarian, S. *Nature* **288**, 613-615 (1980).
15. Zakarian, S. & Smyth, D. G. *Nature* **296**, 250-252 (1982).
16. Smyth, D. G., Massey, D. E., Zakarian, S. & Finnie, M. D. A. *Nature* **279**, 252-254 (1979).
17. Akil, H., Shiomi, H. & Matthews, J. *Science* **227**, 424-426 (1985).
18. Ham, J. & Smyth, D. G. *FEBS Lett.* **175**, 407-411 (1984).
19. Zakarian, S. & Smyth, D. G. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5972-5976 (1979).

ACKNOWLEDGEMENTS. We thank Drs D. G. Smyth, G. Trinchieri and W. H. Wurner for critical review of this manuscript. This work was supported by grants from the NSF (S.Z.) and the NIH (W.K.S.).

Parental origin of mutations of the retinoblastoma gene

Thaddeus P. Dryja*, Shizuo Mukai*, Robert Petersen†, Joyce M. Rapaport*, David Walton*, & David W. Yandell*

Department of Ophthalmology, Harvard Medical School, *Massachusetts Eye and Ear Infirmary, 243 Charles Street, Boston, Massachusetts 02114, and †Children's Hospital, 300 Longwood Avenue, Boston, Massachusetts 02115, USA

RETINOBLASTOMA and osteosarcoma arise from cells that have lost both functional copies of the retinoblastoma gene¹⁻³. Using the cloned retinoblastoma gene⁴⁻⁷ and other linked polymorphic loci, it is possible to reconstruct the sequential loss of the two homologous gene copies that precedes the development of these tumours. In non-hereditary tumours, the loss of each of the two homologues occurs somatically; in hereditary cases, the initial mutation is in the germline. Recently, Toguchida *et al.* reported that the paternally derived copy is preferentially the first one to become mutant during the genesis of non-hereditary osteosarcomas⁸. We report here a similar analysis of patients with retinoblastoma in which we find no such predilection for initial somatic mutations. In contrast, when an initial mutation was a new germline mutation, it was derived from the father, a result which is consistent with new germline mutations arising primarily during spermatogenesis.

The parental origin of the copy of the retinoblastoma gene that initially undergoes mutation before oncogenesis can be determined in two ways. First, if the leukocytes or tumour of a patient have a deletion of one copy of the gene that is detectable by Southern blotting, intragenic restriction fragment length polymorphisms (RFLPs) can be used to deduce the parental origin of the mutant allele. Second, if the tumour is homozygous at polymorphic loci on chromosome 13 where the patient's

leukocytes are heterozygous, it can be reasoned that the set of alleles remaining in the tumour tags the chromosome 13 homologue that suffered the initial mutation. This sequence of events, in which there is a reduction to homozygosity, is common⁹.

We studied 15 consecutive retinoblastomas in which one of these methods was informative. None of the 15 patients had a family history of either retinoblastoma or osteosarcoma. Seven of the patients had unilateral (unifocal) retinoblastoma. As 88–90% of unilaterally affected, simplex patients do not transmit susceptibility to retinoblastoma to their offspring¹⁰, these seven patients were classed as non-hereditary retinoblastoma cases. With regard to the eight patients with bilateral retinoblastoma, it is well established that bilateral cases with a negative family history carry new germline mutations¹⁰.

In four of the 15 tumours, we detected intragenic deletions of the retinoblastoma gene by Southern blotting, using complementary DNA and genomic DNA probes derived from the locus (see Table 1). The deletions found in cases Rb-3, Rb-1 and Rb-145 were initial mutations. In cases Rb-1 and Rb-145 the deletions were new germline mutations, because they were

present in the leukocytes from the patients but not in those of the parents. In case Rb-3 a deletion was a somatic initial mutation, as tumour cells were homozygous for a deletion that was not present in the patient's leukocytes. In one case of bilateral retinoblastoma, Rb-139, the second mutation was a deletion, as tumour cells were heterozygous for a deletion that was absent from the patient's leukocytes.

We deduced the parental origin of the deleted allele in patients Rb-1, Rb-139 and Rb-145 by examining their leukocyte DNA and that from their parents using probes for informative RFLPs that lay within the deleted regions. The deletion found in case Rb-3, as well as the presumed first 'hits' (point mutations or small deletions that escaped detection by the methods used here) in the remaining tumours, could be assigned to one of the parents' chromosome 13 homologues, because the tumour cells were homozygous for a set of alleles within chromosome 13 that was derived from a particular parent (for an example, see Fig. 1).

Among the seven cases of non-hereditary retinoblastoma, the initial mutation occurred on the chromosome derived from the father in three cases, and from the mother in four (see Table 1). According to some models for tumorigenesis, the copy of the retinoblastoma gene that is derived from the father is more susceptible to somatic mutation¹¹. Our data do not support such a model, being more consistent with the null hypothesis, in which there is no differential susceptibility to somatic mutation between the two homologous copies of the gene. Our data conflict with a previous report describing a similar analysis of osteosarcomas⁸, in which the father's copy of the retinoblastoma gene was the first to suffer mutation in 9 out of 10 informative patients. This discrepancy with our data could be a result of the difference in tumour types (osteosarcomas versus retinoblastomas), but it is unclear to us why the paternally derived copy of the retinoblastoma gene should be more vulnerable to somatic mutation in bone as compared with retina.

The eight informative patients with new germline mutations each carried the heritable mutation on the paternally derived chromosome. This result agrees with a previous report describing the parental origin of cytogenetically detectable deletions of chromosome 13 in retinoblastoma patients¹². The deletions were

TABLE 1 Origin of mutations in the retinoblastoma gene

Patient	Sex	Age of father	Age of mother	First hit	Second hit
Somatic initial mutations					
Rb-3	M	NA	NA	Mat.*	Homozyg.
Rb-5	M	29	28	Mat.	Homozyg.
Rb-88	M	25	28	Pat.	Homozyg.
Rb-113	F	26	27	Mat.	Homozyg.
Rb-114	F	22	21	Mat.	Homozyg.
Rb-116	F	29	27	Pat.	Homozyg.
Rb-129	M	38	28	Pat.	Homozyg.
Average age		28.2	26.5		
New germinal initial mutations					
Rb-1	M	28	23	Pat.†	Point mut.
Rb-38	F	26	NA	Pat.	Homozyg.
Rb-101	M	22	21	Pat.	Homozyg.
Rb-104	M	30	25	Pat.	Homozyg.
Rb-111	F	32	30	Pat.	Homozyg.
Rb-132	M	27	26	Pat.	Homozyg.
Rb-139	M	34	26	Pat.	Deletion‡
Rb-145	F	30	21	Pat.§	Point mut.
Average age		28.6	24.6		

Patients are grouped according to whether their tumours were unilateral (somatic initial mutation) or bilateral (new germinal initial mutation). The table gives the age of each parent at the time of birth of the patient. Abbreviations: Pat., paternally derived copy of the retinoblastoma gene; Mat., maternally derived copy; Homozyg., chromosome 13 homozygosity; M, male; F, female; NA, not available; Point mut., a presumed mutation, not detected by Southern blotting. Fragments of primary tumours were recovered from enucleated eyes and frozen within 24 h. Leukocytes were recovered from venous blood samples. DNA was extracted and digested with appropriate restriction endonucleases, and the resulting DNA fragments separated by electrophoresis on agarose gels. DNA was transferred to nitrocellulose or nylon filters and hybridized with radioactively labelled DNA fragments derived from the human retinoblastoma locus (p2R4.7, p123M1.8, p95HS0.5, pH3-8, p7RV1.55L, p88PRO.6, p68RS2.0 and p35RO.6)^{4,18,19} and with DNA fragments that detect other loci on chromosome 13q: D13S1 (p7F12), D13S2 (p9D11), D13S3 (p9A7), D13S4 (p1E8), D13S5 (pHUB8), D13S6 (pHU10) and D13S7 (pHU26) (refs 20 and 21). Alleles were detected by autoradiography. In addition, allelic isozymes at the esterase D locus were analysed by starch gel electrophoresis of homogenates of red cells and tumour cells. One polymorphism within the retinoblastoma gene was analysed using the polymerase chain reaction (see Fig. 1). The parental origin of the initial mutation of the retinoblastoma locus was decided as described in the text.

* Initial mutation was a deletion of 25 kilobases (kb) of the 3' end.

† Initial mutation was a deletion of 135 kb of the 3' end.

‡ Second mutation was a deletion of 42 kb of the 5' end.

§ Initial mutation was a deletion of 20 kb of the 5' end.

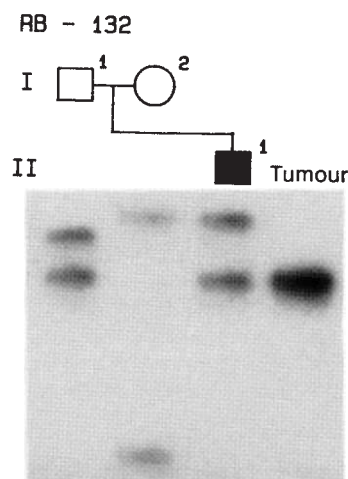


FIG. 1 Paternal origin of the new germline mutation of the retinoblastoma gene in patient Rb-132, who had bilateral retinoblastoma. All four parental copies of the retinoblastoma gene can be distinguished with this polymorphism. Only the paternally derived allele remains in the tumour.

METHODS. The polymorphism assayed here is due to a variable number of tandem repeats of the sequence CTTT(T) that occurs 3' to exon 20 of the retinoblastoma gene. The polymerase chain reaction was used to amplify a 350–400 base-pair sequence containing the repeated sequence. The amplified fragments were radioactively labelled and separated by polyacrylamide gel electrophoresis, as described elsewhere²².

present in peripheral lymphocytes and therefore could be assumed to represent new germline mutations. There were nine cases where the origin of the mutant chromosome could be inferred using chromosome heteromorphisms, and in eight of those cases the mutant chromosome was derived from the father. Combining these data with ours, new germline mutations appeared in the paternally derived chromosome in 16 out of 17 informative cases. This result would be very improbable ($P < 0.0005$) should germinal mutations of the retinoblastoma gene be equally likely in maternally and paternally derived chromosomes.

The paternal origin of most new germline mutations is easily explained if they arise predominantly during spermatogenesis. Such a hypothesis would be supported by a paternal age effect, in which new cases of bilateral retinoblastoma might occur more frequently in the offspring of older fathers. Although we detected no paternal age effect in our small set of patients (see Table 1), a review of published data allowed Vogel and Rathenberg to conclude that a weak paternal age effect does exist for bilateral retinoblastoma¹³. Furthermore, if germline mutations arise predominantly during DNA replication, then the observed excess of germline mutations in paternal versus maternal copies of chromosome 13 (16:1) should correlate with the number of cell divisions from embryonic development to meiosis in males versus females. Vogel and Rathenberg provide the following estimates: 380 divisions for sperm from a 28-year-old male; 23 for ova from an adult female. The ratio (380:23 or 16.5:1) agrees with the data.

In summary, our investigations of the parental origin of mutations of the retinoblastoma gene cast serious doubt on the notion that initial somatic mutations occur predominantly on the chromosome 13 derived from the father, and they eliminate a theoretical requirement for genomic imprinting as a factor in somatic mutagenesis at the retinoblastoma locus. A comparison of our results with studies of chromosome 11p in Wilms tumour^{14,15}, where the paternally derived chromosome was preferentially retained in tumour cells, is complicated by the fact that the locus governing hereditary susceptibility to Wilms tumour is not on that chromosome arm^{16,17}. Furthermore, those studies included patients with bilateral disease, indicating a constitutional, if not heritable, susceptibility to the tumours. On the other hand, our results and those from a previous study¹² convincingly demonstrate that the majority of new germline mutations of the retinoblastoma gene arise in the paternal chromosome 13. Future investigations of the high mutation rate of the retinoblastoma locus should explore mutagenic factors peculiar to fathers. □

Received 29 March; accepted 11 May 1989.

- Knudson, A. G. Jr *Proc. natn. Acad. Sci. U.S.A.* **68**, 820-823 (1971).
- Comings, D. E. *Proc. natn. Acad. Sci. U.S.A.* **70**, 3324-3328 (1973).
- Friend, S. H., Dryja, T. P. & Weinberg, R. A. *New Engl. J. Med.* **318**, 618-622 (1988).
- Dryja, T. P., Rapaport, J. M., Joyce, J. M. & Petersen, R. A. *Proc. natn. Acad. Sci. U.S.A.* **83**, 7391-7394 (1986).
- Friend, S. H. *et al. Nature* **323**, 643-646 (1986).
- Lee, W.-H. *et al. Science* **235**, 1394-1399 (1987).
- Fung, Y.-K. *et al. Science* **236**, 1657-1661 (1987).
- Toguchida, J. *et al. Nature* **338**, 156-158 (1989).
- Dryja, T. P. *et al. New Engl. J. Med.* **310**, 550-553 (1984).
- Vogel, F. *Hum. Genet.* **52**, 1-54 (1979).
- Reik, W. & Surani, M. A. *Nature* **338**, 112-113 (1989).
- Ejima, Y., Sasaki, M. S., Kaneko, A. & Tancoka, H. *Hum. Genet.* **79**, 118-123 (1988).
- Vogel, F. & Rathenberg, R. *Adv. hum. Genet.* **5**, 223-318 (1975).
- Schroeder, W. T. *et al. Am. J. hum. Genet.* **40**, 413-420 (1987).
- Mannens, M. *et al. Hum. Genet.* **81**, 41-48 (1988).
- Grundy, P. *et al. Nature* **336**, 374-376 (1988).
- Huff, V. *et al. Nature* **336**, 377-378 (1988).
- Wiggs, J. *et al. New Engl. J. Med.* **318**, 151-157 (1988).
- Bookstein, R. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 2210-2214 (1988).
- Dryja, T. P., Rapaport, J. M., Weichselbaum, R. & Bruns, G. A. P. *Hum. Genet.* **65**, 320-324 (1984).
- Cavenee, W., Leach, R., Monandus, T., Pearson, P. & White, R. *Am. J. hum. Genet.* **36**, 10-24 (1984).
- Yandell, D. W. & Dryja, T. P. *Am. J. hum. Genet.*, in the press.

ACKNOWLEDGEMENTS. We thank Tracey Campbell, Siri Dayton and Terri McGee for Fig. 1. This research was supported by the NIH and the American Cancer Society, by the Alcon Research Institute Inc., and by gifts to the Taylor R. Smith Laboratory. T.P.D. is a Research to Prevent Blindness-Dolly Green Scholar.

The *Neurospora* clock gene frequency shares a sequence element with the *Drosophila* clock gene period

C. Robertson McClung*, Barbara A. Fox & Jay C. Dunlap†

Department of Biochemistry, Dartmouth Medical School, Hanover, New Hampshire 03756, USA

THE isolation and characterization of single gene mutations affecting the circadian biological clocks of several organisms (reviewed in ref. 1) has left little doubt that circadian rhythms can be subjected to classical genetical analysis. Many of these mutations occur at the same few genetic loci (*frequency* (*frq*) in the fungus *Neurospora*², and *period* (*per*) in fruit fly *Drosophila*³); these loci represent the best studied clock-affecting genes known. Mutant strains are usually affected in more than one basic clock property^{1,4}, suggesting an inter-relatedness at the molecular level among these basic properties that would not have been predicted *a priori*. The extensive background information available concerning the *frq* locus⁴ provides a basis for the molecular dissection of the *Neurospora* circadian clock—the most minimal circadian system thus far described. We report here the cloning and analysis of the *frq* locus and show it to be larger and more complex than would have been predicted from the available genetic data. Complete rescue of all of the pleiotropic mutant phenotypes^{5,6} of the recessive *frq*⁹ allele requires transformation with a 7.7-kilobase (kb) region of DNA encoding at least two transcripts. Sequence analysis of this region has allowed the identification of a common element between *frq* and *per* which, given the background similarities in their classical genetic characteristics, suggests the possibility of a common element in the clock mechanisms of these two organisms.

The *frq* locus of *Neurospora crassa* lies on the right arm of linkage group VII (VII R), ~2 map units distal to *oligomycin resistance* (*oli*) and 2.5 map units proximal to *formate* (*for*) (Fig. 1a). A chromosomal walk was undertaken starting from the *oli* gene⁷ and covering ~190 kb corresponding to ~8 map units⁸ along VII R (Fig. 1b). Individual phage and cosmids arising from the walk were assayed by transformation⁹ for their ability to rescue the circadian banding pattern of the recessive *frq*⁹ allele (Fig. 1b). Four partially overlapping cosmids arising from the right-hand side of the walk were capable of rescuing the *frq*⁹ phenotype, providing strong presumptive evidence that these four cosmids each contained *frq*. This was confirmed in two ways. First, in agreement with expectations based on the genetic map, the next rightward cosmid (31:5E) was found to be capable of rescuing *for* mutations (Fig. 1). Second, DNA fragments arising from cosmids 2:10A, 8:3B and 23:9D were mapped by restriction fragment length polymorphism analysis¹⁰, confirming for all three a genetic map location on VII R (1/38 recombinants with *ars-1* near the centromere, 6/38 with *nic-3* on VII L; data not shown). We have thus established a physical map covering ~8 map units along the right arm of linkage group VII, shown that the physical map corresponds well to the known genetic map, and identified both the *frq* and *for* genes.

At this resolution the *frq* locus was defined by the 13 kb that span the region of overlap of the four *frq*⁹ complementing cosmids. This region was subjected to restriction analysis, subcloning, and retransformation of subclones into *frq*⁹ (Fig. 1c). Based on previous experience with the *frq*-containing cosmid clones, *for*, and other genes, two outcomes of these transforma-

* Present address: Department of Biological Sciences, Dartmouth College, Hanover, New Hampshire 03755 USA.

† To whom correspondence should be addressed.

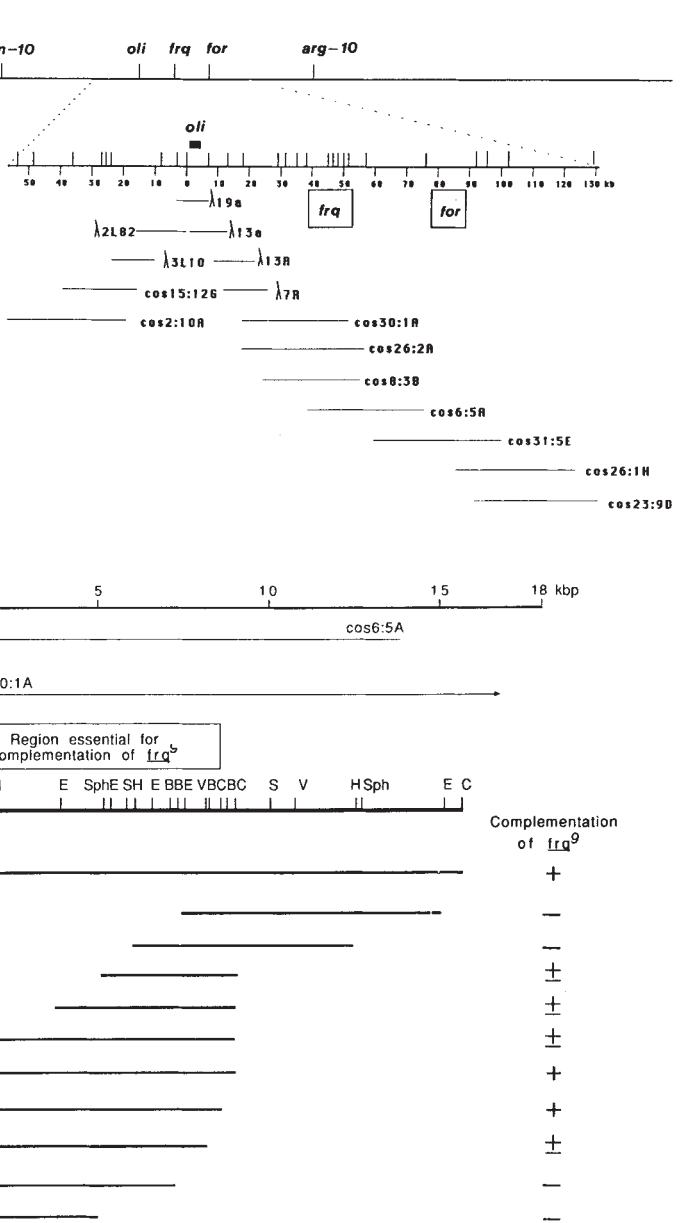
tion/rescue experiments were anticipated: either a complete complementation of the pleiotropic *frq*⁹ mutation or a complete lack thereof. Surprisingly there appeared a third intermediate phenotype (denoted $\pm$) in which some of the mutant characteristics of *frq*⁹ were complemented (see Fig 1c, legend) but in which there was still an absence of circadian conidial banding; the reverse phenomenon (complementation of circadian banding in the absence of robust conidiation) was not observed. These studies have shown that the smallest region of DNA capable of complete complementation of the circadian regulation phenotype is the ~ 7.7 kb insert of pKAJ101. In the transformation assay, *frq* generally displays a variable phenotype in that the circadian banding pattern of *frq*⁹ strains transformed to benomyl resistance and containing *frq*⁺ DNA is phenotypically transformed only $\sim 20\%$ of the time. But in nearly all cases where we have examined this subset of transformed strains, the

period length of the circadian oscillation has been indistinguishable from wild-type (average period length among transformants 20.9 ± 1.8 h (s.d.), $n = 60$ versus 21.6 ± 0.5 h, $n = 5$ for wild-type), and the pattern of temperature compensation appears to be completely normal (Fig. 2). There is a suggestion, however, that the variability of the period length was consistently greater among transformants than in wild-type strains. Transformants could be entrained normally to daily 5-minute light pulses but in some cases the stable phase of entrainment (that is, the time of day at which the centre of the conidial band occurred) was different from wild-type.

The *frequency* locus gives rise to at least two processed transcripts, one ~ 1.5 kb in size arising from the left-hand side of the region and one or more large overlapping transcripts of ~ 5 kb arising from the right-hand side (Fig. 3). The origin of these transcripts within the *frq*-complementing region, as

FIG. 1 Localization of *frq* gene sequences on *Neurospora crassa* linkage group (LG) VIII: correlation of the genetic map and the physical map via chromosomal walking from *oli*. **a**, Genetic map of LG VIII, with the centromere indicated by the open circle at left. Genetic distances are *oli*-*frq* ~ 2 map units and *frq*-*for* ~ 2 map units. **b**, Physical map of the region of LG VIII around *oli* (solid box). Dotted lines extending from the genetic map (**a**) to the physical map indicate the approximate correlation of the two maps. Vertical lines above the map indicate *EcoRI* restriction sites; restriction sites for other enzymes are not shown for the sake of clarity. Lines below the map indicate the positions of inserts of phage and cosmid clones identified in the walk. Open boxes beneath the physical map indicate the initial positioning of *frq* and *for* genes. **c**, Expansion of the region defined as common to the four cosmids complementing *frq*⁹ (open box labelled *frq* in **b**). Restriction enzymes are *Bam*HI (B), *Cl*al (C), *Eco*RI (E), *Eco*RV (V), *Hind*III (H), *S*alI (S) and *Sph*I (Sph). Inserts of plasmid subclones used to transform *frq*⁹ recipients are indicated by the narrow lines below the restriction map. Insert end points correspond to restriction sites indicated in the restriction map with the exception of the right ends of pCRM101, pCRM102, pCRM105, pCRM108 and pKAJ101 which are defined by the insert-vector junction of cos 30:1A. The ability of the subclones to rescue the *frq*⁹ phenotype is indicated at the right; - denotes no phenotypic rescue, + denotes full restoration of circadian banding with wild-type period length, and restoration of wild-type levels of pigmentation and conidiation, and $\pm$ denotes restoration of wild-type pigmentation and conidiation but no restoration of circadian banding (see text).

METHODS. Steps were taken in a λ genomic library (M. Orbach and C. Yanofsky, Stanford University) and later in an ordered cosmid library⁹; chromosomal sequences between cos 15:12G and cos 30:1A are not represented in the Vollmer-Yanofsky cosmid library. Cosmids and subclones²⁴ were used to transform *frq*⁹ *Neurospora* to benomyl resistance (Bm^r); phage clones were co-transformed with pSV50 (ref. 9). Primary transformants (~ 500 per μ g DNA) were identified after three days growth at 30 °C, transferred to Horowitz Complete²⁵ slants supplemented with 500 ng ml⁻¹ benomyl (Bm), initially screened for phenotypic rescue by visual inspection of conidial pigmentation and subsequently screened on race tubes⁶. Media for race tubes of Bm^r transformants did not contain Bm because the presence of as little as 100 ng ml⁻¹ of the drug interfered with conidial banding pattern even in resistant strains (data not shown). For the identification of *for* containing cosmids and subclones, primary transformants of *for*⁻ were identified on Horowitz Complete medium plates supplemented with 500 ng ml⁻¹ Bm and 500 μ g ml⁻¹ sodium formate, and



individual transformants then screened in Vogel's minimal liquid cultures with or without sodium formate. The *for* gene was placed between 74 and 77 kb distal from *oli* by the finding that the *for*-complementing ability unique to 31:5E was preserved following digestion with *Hind*III and *Bgl*II (data not shown).

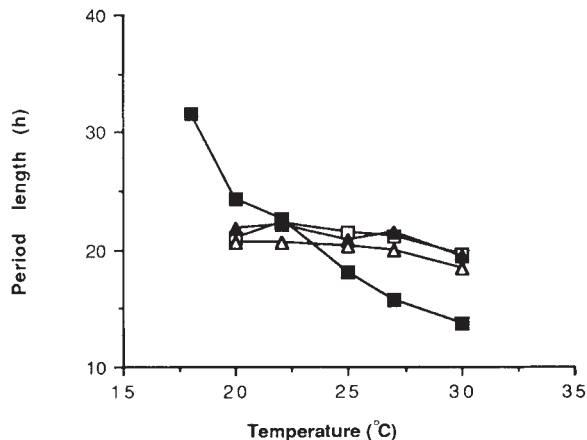


FIG. 2 Period length and temperature compensation characteristics of the clock are completely restored in phenotypically rescued *frq*⁹ transformants. Period length of the rhythm is plotted versus temperature for *bd,frq*⁺ (open squares), *bd,frq*⁹ (closed squares) and for *bd,frq*⁹ strains transformed with pCRM101 (closed triangles) or pCRM102 (open triangles). For comparison, *bd,frq*⁹ data are replotted from ref. 5.

METHODS. A single race tube having six to eight circadian cycles was considered to yield a single estimate of period length²⁶. Points here were the average of between three and 25 race tubes, depending on the experiment, and standard deviations, omitted for the sake of clarity, ranged from 0.3 h to 1.5 h; error bars at ± 1 s.d. overlap for all points except *frq*⁹.

determined by hybridization analysis using different subcloned pieces (Fig. 3), agrees well with the limits of the *frq* region as determined by phenotypic analysis of transformants receiving different regions of DNA (Fig. 1c). It thus appears that DNA sufficient to allow expression of all of the transcripts must be present to rescue fully *frq*⁹ recipients, whereas to a first approximation, clones containing only the region encoding the long transcript(s) are sufficient for partial rescue.

Figure 4a contains the DNA sequence for an 8,657 base-pair (bp) region of genomic DNA containing the inserts of pKAJ101 and pCRM101, the interval containing the transcription unit(s) encoding the *frq* gene products. Within the first (5') 200 bp of pKAJ101 (that is, within the region shown to be required for *frq* complementation; compare pCRM116 with pKAJ101), there exists an extended inverted repeat (18/22 matches separated by 100 bp of potential loop) which if transcribed could yield a stable stem loop ($\Delta G(25^\circ\text{C}) = -23.7$ kcal; ref. 11) of the type implicated in translational control under some conditions¹². Computer analysis of the entire sequence 3' of this reveals the presence of 11 open reading frames (ORFs) greater than 100

amino acids in length, one of which deserves special comment. Within the region corresponding to the long transcript(s), there is a single very long ORF of 2,364 bp that extends from nucleotide 5174 to 7538 and appears in only one of the three possible reading frames (Fig. 4). Predicted codon usage within this region, based on a conceptual translation, is strongly biased in a manner typical for expressed genes in *Neurospora*¹³. This would not be expected in randomly chosen nucleotide triplets and suggests that at least part of this ORF may be expressed and can be used to predict a part of the amino-acid sequence of the *frq* gene product (see below).

The entire *frq* sequence shown in Fig. 4a (as four overlapping 3,000-nucleotide windows) was used to search GENBANK (July, 1988 release) using FASTA¹⁴. These searches failed to identify any gene or protein with extended homology to *frq* and, with one exception, none of the genes or gene products identified suggested any biologically relevant relationship. The region of DNA extending from nucleotides 4001 to 7000, however, did identify a sequence element in the *per* gene of *Drosophila* similar (*Z* value of 3.5, ref. 15) to a region within *frq* (extending from nucleotides 5776 to 5830). The long ORF referred to above lies within this segment of DNA. Although this similarity would be classified only as possibly statistically significant¹⁶, we find it noteworthy for the following reasons: (1) Although several other sequences bearing weak similarities to *frq* were detected, these were generally based on small regions of highly biased nucleotide composition and were of marginal significance; (2) The similarity with *per* apparently extends beyond the nucleic acid to the protein sequence. When the conceptual translation of the region of similarity within the long ORF was used to search the NBRF (National Biomedical Research Foundation) database using FASTA¹⁴, the homology to *per* reappeared, again among the strongest of a number of weak similarities that included several proteoglycans. Part, but not all, of the region of *per* sharing similarity with the *frq* long ORF contains a repeating unit of threonine and glycine (TG repeat^{17,18}). This region is contained within the longest exon in the *per* transcription unit; (3) part of the region of similarity in the *frq* long ORF is a modified TG repeat (Fig. 4b) where there is considerable diversity in the third position among the codons used. Thus it does not appear simply as an oligonucleotide repeat. Furthermore, the similarity between *frq* and *per* is not confined to the direct TG repeat; it extends both 3' and 5' beyond the direct repeat region of both genes (Fig. 4b).

The region containing the TG repeat is one of the salient features of the *per* gene product in *Drosophila*, where it may act as the site for some of the extensive glycosylation thought to be present in the *per* protein^{17,19}. As in *per*, the region in the *frq* long ORF within and around the similarity region contains several possible sites for both N-linked and O-linked

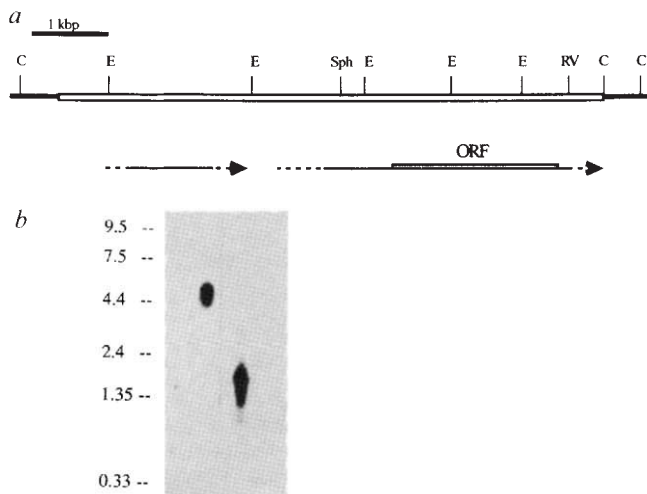


FIG. 3 Transcription of the *frq* locus region. *a*, Map of the region required for rescue of the *frq*⁹ phenotype. Sites are as in Fig. 1c. Arrows beneath the map indicate transcripts detected in poly(A)⁺ RNA from *bd,frq*⁺ mycelia; open box above the long transcript indicates the position of the long ORF (see text). Dotted lines indicate imprecision concerning transcriptional initiation and termination sites. *b*, RNA blot analysis of transcripts prepared from *bd,frq*⁺ mycelia. Left hand panel shows hybridization to the *ClaI/SphI* fragment extending from nucleotide 4,604 to 8,180 in Fig. 4a; right hand panel shows hybridization to the *EcoRI* fragment extending from 1,201 to 3,207.

METHODS. Blots using RNA isolated from *bd,frq*⁺ mycelia were prepared and processed for autoradiography as described previously²⁷. Direction of transcription and localization of the transcripts were determined using single-stranded DNA probes²⁸ generated from each of the *EcoRI* fragments shown in the restriction map in *a*.

FIG. 4 a, DNA sequence of the *frq* gene region. Position 1 of the sequence corresponds to the *Clal* site at the left hand of the restriction map in Fig. 1c. The location of an inverted repeat near the beginning of the sequence (see text) is noted by arrows below the sequence and the end of the insert of cosmid 30:1A is noted by an asterisk above the sequence at position 442. Deletion 3' of this to the *EcoRI* site (underlined) at position 1,202 (pCRM116) results in fragments capable of only partial rescue. The *Clal* site (underlined) at position 8,179 marks the end of pKAJ101. Deletion 5' of this to the *EcoRV* site (underlined) as position 7,765 (pCRM105) results in fragments capable of only partial rescue. Deletion further 5' to the *BamHI* site (underlined) at position 6,619 (pBAM4.1) completely eliminates all phenotypic rescue (see text and Fig. 1c). Conceptual translation product of a 788-amino-acid polypeptide encoded by the open reading frame starting at position 5,174 is indicated below the DNA sequence, and the region bearing strongest similarity to the *per* gene is underlined. The numbers in the right margin indicate nucleotide number from the beginning. b, Alignment of the core of the region of similarity between the conceptual translations of *frq* and *per*. Amino-acid identities are boxed and residues generally described as conservative replacements^{29,30} are denoted by a dot.

METHODS. a, Ordered sets of overlapping deletion subclones³¹ were generated from *Neurospora* genomic DNA cloned into the Bluescript series of vectors (Stratagene), and gaps in the sequence were resolved by cloning specific subregions. The nucleotide sequence was determined for both strands by the dideoxy method³². b, Alignment was generated using the Align program³³ as implemented in the FASTA program package¹⁴.

b *frq* S S K D N G S A S N S G G D Q T E L G G T G T G S G D G S S G G R T G N N T S P P G A I A P D
per S S V T N T S I A G T G G T G T G (T G) T G T G T G T G T G T G T G T G N T S S G T G T G T A

glycosylation, thus suggesting that the *frq* gene product could, like *per*, be glycosylated. Additionally, analysis of the conceptual protein sequence both 5' and 3' of the similarity region reveals the presence of several potential kinase A and kinase C phosphorylation sites²⁰, although in the absence of a complete complementary DNA sequence, speculation as to the potential importance of these sites seems premature.

The apparent size and diversity of the *frq* products are noteworthy, particularly in light of previous genetic studies. Fine structure mapping of the existing *frq* alleles⁶ has suggested that they lie within ~0.003 map units of each other, a region comprising a few as 100 bp. Preliminary studies on the *frq* mutant alleles have identified no differences on Southern blots (data not shown), thus eliminating the presence of gross structural distortions in the DNA (which might reduce recombination frequencies) as the molecular basis of these mutations. We suggest therefore that there may be a region of *frq* which is either particularly mutable or particularly important to its function as a part of the circadian clock.

In addition to the surprising size and molecular complexity of the *frq* locus, the identification of a region of similarity with *per* was unexpected. In selecting *per*, the search identified the only other clock gene that has been cloned to date, a gene whose mutations confer phenotypes similar in many ways to the documented phenotypes of the various *frq* alleles^{1,2}. The sequence similarity is not extensive so that its existence may be merely coincidental, and given our present state of knowledge of the products of *frq* it is clearly impossible to conclude that the two genes share a common ancestor. Alternatively it has been something over 1×10^9 years since plants and fungi diverged from animals during the Precambrian era²¹, plenty of time for genetic drift to eliminate similarities. Because a part of the region of similarity in each gene contains similar structural features including several potential glycosylation sites, this similarity could reflect convergent evolution to protein structural elements necessary for the function of these and perhaps other clock proteins. Indeed, if the similarity is biologically relevant, it is noteworthy and suggests that 'clock' genes in other organisms might be identified through hybridization analysis. In light of this, the identification of genes in other organisms based on hybridization to *per*^{22,23} may take on a renewed significance. In hybridization experiments at moderate stringency, DNA from the *frq* locus has been used to screen genomic DNAs from a number of vertebrate, invertebrate, fungal and plant systems. The identification in this way of potential homologues in yeast, *Arabidopsis* (J.C.D., S. Kay, A. Miller, N.-H. Chua; unpublished data), soybean and several mammalian species, combined with the novel ability to transform and genetically manipulate these organisms, suggests the possibility of extending the molecular genetic analysis of circadian clocks to these species as well. □

21. Dayhoff, M. O. in *Atlas of Protein Sequence and Structure* Vol. 5, suppl. 3. (ed. Dayhoff, M. O.) 1-8 (Natl. biomed. Res. Found., Silver Spring, 1978).
22. Shin, H. S., Bargiello, T. A., Clark, B. T., Jackson, F. R. & Young, M. W. *Nature* **317**, 445-448 (1985).
23. Weber, M. L., deGroot, E. J. & Schweiger, H. G. *Molec. gen. Genet.* **209**, 1-7 (1987).
24. McClung, R. M., Phillips, J. D., Orbach, M. J. & Dunlap, J. C. *Expl. Mycol.* (in the press).
25. Davis, R. L. & deSerres *Meth. Enzym.* **71A**, 79-143 (1970).
26. Dunlap, J. C. & Feldman, J. F. *Proc. Natn. Acad. Sci. U.S.A.* **85**, 1096-1100 (1988).
27. Loros, J. J., Denome, S. & Dunlap, J. C. *Science* **243**, 385-388 (1989).
28. Hu, N. & Messing, J. *Gene* **17**, 271-277 (1982).
29. Feng, D. F., Johnson, M. S. & Doolittle, R. F. *J. molec. Evol.* **21**, 1112-125 (1985).
30. Schwartz, R. M. & Dayhoff, M. O. in *Atlas of Protein Sequence and Structure* Vol. 5, Suppl. 3 (ed. Dayhoff, M. O.) 353-358 (Natl. biomed. Res. Found., Silver Spring, 1978).
31. Henikoff, S. *Gene* **28**, 351-359 (1984).
32. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
33. Myers, E. & Miller, W. *CABIOS* **4**, 11-17 (1988).

ACKNOWLEDGEMENTS. We thank J. Phillips and S. Denome for technical assistance, Jennifer Loros and B. Masucco for help in analysis of race tubes, S. Brody for pointing out potential phosphorylation sites, and J. Loros, J. Hall and J. Feldman for comments on the manuscript. This work was supported by the NIH and the Norris Cotton Cancer Center (J.C.D.).

RNA helicase activity associated with the human p68 protein

H. Hirling, M. Scheffner, T. Restle & H. Stahl

Fakultät für Biologie, Universität Konstanz, D-7750 Konstanz, FRG

IT has been proposed that p68, a nuclear protein of relative molecular mass 68,000, functions in the regulation of cell growth and division¹. A complementary DNA analysis of the protein² has revealed extensive amino-acid sequence homology to the products of a set of genes recently identified in organisms as diverse as *Escherichia coli* and man, which include the eukaryotic translation initiation factor eIF-4A (refs 3-9). The protein products of the new gene family have several motifs in common which are thought to be involved in nucleic acid unwinding¹⁰⁻¹². As yet, however, only eIF-4A, through its effect on RNA, has been shown to possess unwinding activity^{13,14}. Here we report that purified p68 also exhibits RNA-dependent ATPase activity and functions as an RNA helicase *in vitro*. The protein was first identified by its specific immunological cross reaction with the simian virus 40 large T antigen¹, the transforming protein of a small DNA tumour virus¹⁵. Surprisingly, T antigen also has an RNA-unwinding activity¹⁶: the homology between the two polypeptides, although confined to only a small region resembling the epitope of the cross-reacting antibody (PAb204) (ref. 2), should therefore be of functional significance. Furthermore, the RNA-unwinding activity may be involved in the growth-regulating functions of both proteins.

Cell fractionation analysis has revealed a close association of the p68 protein with the nuclear matrix (unpublished data). Therefore, conditions were developed to quantitatively extract native p68 from cell nuclei, followed by a conventional chromatographic purification program (see Fig. 1a legend for details). Essentially homogenous p68 protein was obtained which sedimented at 4-5 S in a sucrose gradient (Fig. 2b).

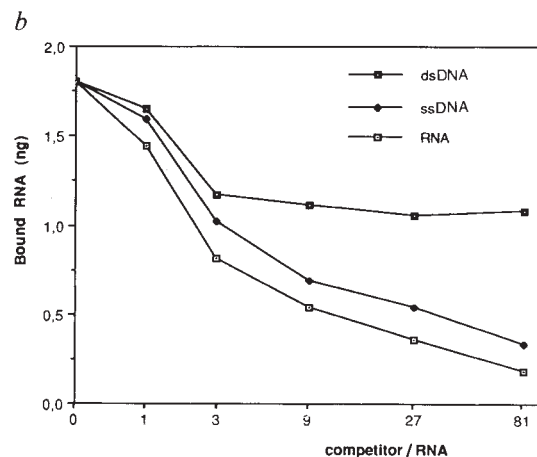
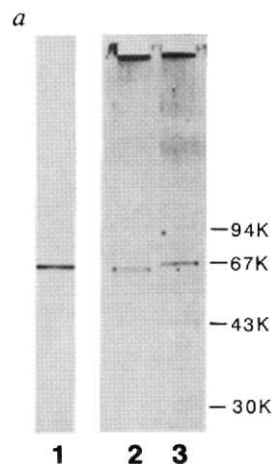
During protein purification we made use of a single-stranded (ss) DNA cellulose column to which p68 bound with high affinity. The binding of purified p68 to nucleic acid was further tested using radiolabelled RNA in a nitrocellulose filter assay. Efficient RNA binding (resistant to salt up to 250 mM; data not shown) was observed irrespective of the presence or absence of Mg²⁺ and ATP, and could be best competed with RNA (total cytoplasmic RNA from HeLa cells) or ss M13 DNA (Fig. 1b). Double-stranded (ds) DNA competed only partially for RNA binding, and no binding of p68 to ds DNA was found at 100 mM NaCl (data not shown). RNA binding has also been reported for the *E. coli* SrmB protein, another member of the eIF-4A-like family, and this activity may be a function of the relatively basic C-terminus of the two proteins⁵. Factor eIF-4A lacks this region and the smaller eIF-4A protein binds RNA only in the presence of a second protein¹⁷.

The SrmB protein, like eIF-4A, is an RNA-dependent

Received 22 February; accepted 8 May 1989.

1. Hall, J. C. & Rosbash, M. A. *Rev. Neurosci.* **11**, 373-393 (1988).
2. Gardner, G. F. & Feldman, J. F. *Genetics* **96**, 877-886 (1980).
3. Konopka, R. J. *Fed. Proc.* **38**, 2602-2605 (1979).
4. Feldman, J. F. & Dunlap, J. C. *Photochem. Photobiol. Rev.* **7**, 319-368 (1983).
5. Loros, J. J. & Feldman, J. F. *J. Biol. Rhythms* **1**, 187-198 (1986).
6. Loros, J. J., Richman, A. R. & Feldman, J. F. *Genetics* **14**, 1095-1110 (1986).
7. Viebrock, A., Perz, A. & Siebald, W. A. *EMBO J.* **1**, 565-571 (1982).
8. Perkins, D. D., Radford, A., Newmeyer, D. & Bjorkman, M. *Microbiol. Rev.* **46**, 426-570 (1982).
9. Vollmer, S. J. & Yanofsky, C. *Proc. natn. Acad. Sci. U.S.A.* **83**, 1-5 (1986).
10. Metzberg, R. L., Stevens, J. N., Selker, E. U. & Morzycka-Wroblewska, E. *Proc. natn. Acad. Sci. U.S.A.* **82**, 2067-2071 (1985).
11. Tinoco, I. et al. *Nature New Biol.* **246**, 40-41 (1973).
12. Kozak, M. *Molec. cell. Biol.* **8**, 2737-2744 (1988).
13. Orbach, M. J., Porro, E. A. & Yanofsky, C. *Molec. cell. Biol.* **6**, 2452-2461 (1986).
14. Pearson, W. R. & Lipman, D. J. *Proc. natn. Acad. Sci. U.S.A.* **85**, 2444-2448 (1988).
15. Lipman, W. J. & Pearson, W. R. *Science* **227**, 1435-1441 (1985).
16. Doolittle, R. F. *Of URFs and ORFs* (University Science Books, Mill Valley, 1986).
17. Jackson, F. R., Bargiello, T. A., Yun, S. H. & Young, M. W. *Nature* **320**, 185-188 (1986).
18. Citri, Y. et al. *Nature* **326**, 42-47 (1987).
19. Reddy, P., Jacquier, A. C., Abovich, N., Petersen, G. & Rosbash, M. *Cell* **46**, 53-61 (1986).
20. Gonzalez, G. A. et al. *Nature* **337**, 749-752 (1989).

FIG. 1 *a*, SDS-PAGE of the purified p68 protein. *b*, RNA-binding activity of purified p68. During purification, p68 was traced by immunoblotting using PAb204 (ref. 1). Growing human 293 cells (5×10^8) were lysed with NP40, and nuclei were extracted with 25 mM Tris-HCl, pH 10.5; 1 mM EDTA; 0.5 M NaCl; 10 mM β -mercaptoethanol; 0.5% Triton X100. The S100 supernatant fraction (buffered with Tris-HCl, pH 8.5) was adsorbed to hydroxyapatite. Proteins were eluted with potassium phosphate (pH 8.0, 350 mM) and precipitated with ammonium sulphate (30% saturation). The protein precipitates were washed in A buffer (25 mM Tris-HCl, pH 8.5; 2 mM EDTA; 10 mM β -mercaptoethanol; 7.5% glycerol) and subsequently dissolved in A buffer containing 300 mM ammonium sulphate and 50 mM EDTA. The protein solution (diluted sixfold with A buffer) was fractionated by chromatography on a 1 ml MonoQ column with a Pharmacia FPLC system. Following elution with a 0.05–1 M gradient of NaCl in A buffer, p68 was identified in fractions containing 500–700 mM salt. Pooled fractions were diluted with A buffer (1:6, vol/vol) and chromatographed on a ssDNA-cellulose column (2 ml) washed with A buffer plus 250 mM NaCl. Proteins eluting at 700 mM NaCl were concentrated by a second ammonium sulphate precipitation. Suspended protein was sedimented through a sucrose gradient (0–15% sucrose in 25 mM Tris-HCl, pH 7.8; 20% glycerol; 1 mM EDTA; 50 mM NaCl; 10 mM β -mercaptoethanol; 0.1% Triton X100) at 48,000 r.p.m. and 0 °C in a Beckman SW55 rotor for 20 h and fractions containing p68 ($\sim 10 \mu\text{g}$) were pooled and stored at -70 °C. Analysis of an aliquot of the pooled fractions (representing 150 ng protein) by SDS-PAGE and silver staining is shown (lane 2). Lane 1 shows



a western blot analysis of an identical probe using monoclonal PAb204. Relative molecular masses (in thousands) of standard proteins are shown on the right (lane 3). Binding of p68 to RNA was determined by the retention of [^{32}P]RNA (synthesized by run-off transcription of the *PvuII*-restricted vector pGEM3, Promega) on a nitrocellulose filter as previously described⁶. Reaction mixtures contained 4 ng RNA (10^4 d.p.m. ng^{-1}) plus 10 ng p68 and were incubated for 30 min on ice in the presence or absence of competing nucleic acids (dsDNA and ssDNA were from M13; RNA, total cytoplasmic RNA from HeLa cells).

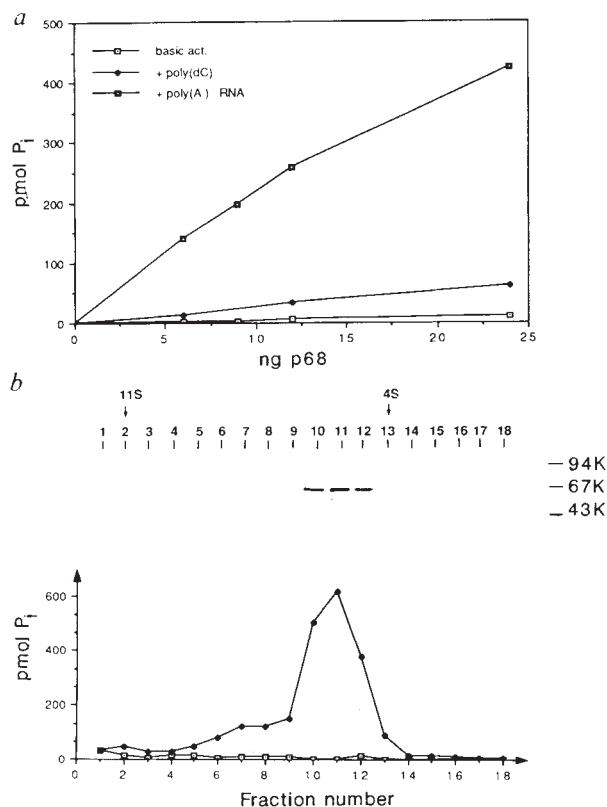


FIG. 2 ATPase activity of p68. *a*, Dependence on nucleic acids. ATPase activity was determined at 0.1 mM [γ - ^{32}P]ATP (2×10^5 d.p.m. nmol^{-1}) as previously described¹⁸. Where indicated, reaction mixtures (50 μl) contained 4 ng poly(dC) or the same amount of poly(A)⁺ RNA from HeLa cells. *b*, Co-sedimentation of the ATPase activity with p68. Purified p68 was sedimented in a sucrose gradient (see legend to Fig. 1*a*). The fractions were analysed for p68 by western blotting (upper, with M_r markers) and for ATPase activity (lower part) in the presence or absence of poly(A)⁺ RNA as described above. Sedimentation was from right to left. The positions of the sedimentation markers (4S, haemoglobin and 11S, catalase) were determined in a parallel gradient.

ATPase⁶. We found that this is also a function performed by purified p68 protein (Fig. 2). Significant ATP hydrolysis was observed in the presence of poly(C), a synthetic RNA, or poly(A)⁺ RNA, but almost no activity was detected in the absence of nucleic acids, and ssDNA had little stimulatory effect. The ATPase copurified with p68 and cosedimented during sucrose gradient centrifugation (Fig. 2*b*). The PAb204 antibody had no influence on the p68 ATPase, which was not surprising as the antibody epitope lies within the C-terminus and is remote from the putative helicase consensus motifs of the eIF-4A-like family². The specific ATPase activity of our p68 preparation ($11 \text{ pmol Pi s}^{-1} \mu\text{g}^{-1}$) is one or two orders of magnitude higher than that described for the SrmB protein⁶ or the eIF-4A (ref. 18), respectively.

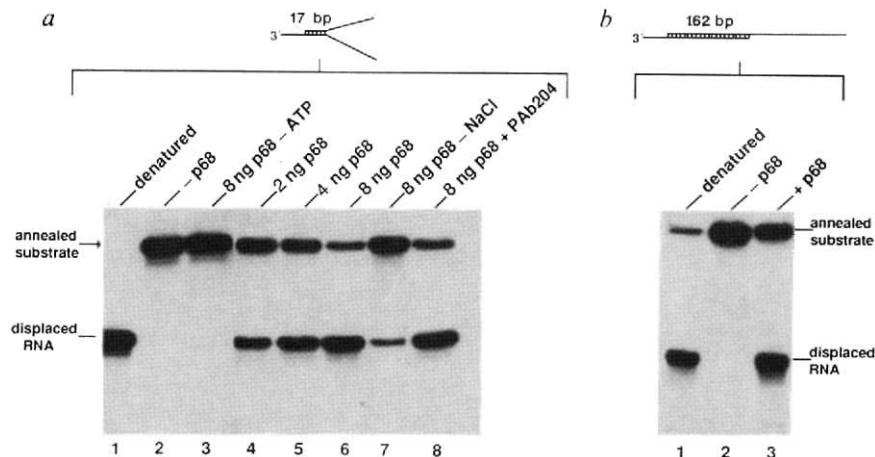
To test for an RNA unwinding activity, we used partially dsRNA substrates (Fig. 3) and, irrespective of their length, found that stretches of these substrates were indeed unwound. The unwinding reaction was strongly stimulated by salt (100 mM; compare lanes 4, 5 and 6 with lane 7) and was dependent on p68 concentration (lanes 4–6), ATP (lane 3), Mg^{2+} and incubation time (data not shown).

Although it seems that RNA-unwinding ability is not *a priori* inherent in DNA helicases, we have recently shown that the SV40 T antigen, which is known to be an ATP-dependent DNA helicase¹⁹, also functions to unwind RNA¹⁶. In spite of the differences between p68 and T antigen, the fact that both have RNA helicase function implies therefore that the short region of limited amino-acid sequence homology between the proteins, the PAb204 binding site, is not fortuitous. As the epitope is not among the consensus motifs defined for the putative helicase family^{2,12}, it may represent an additional function common only to a subset of the gene family.

Although the human p68 was proposed by others to be a DNA helicase *in vitro*², we failed to detect any DNA-unwinding activity using several DNA substrates of different structure (ref. 20 and data not shown). This finding is consistent with the results reported above: the ATPase activity of p68 was not appreciably stimulated by ssDNA, a property essential for DNA helicases. Instead, we observed that p68 efficiently inhibited DNA unwinding by T antigen and the *E. coli* DNA helicases I

FIG. 3 RNA unwinding activity of p68. Analysis of the unwinding reaction by SDS-PAGE and autoradiography. The RNA substrates are schematically drawn above (the lower RNA strand was radioactively labelled in both cases). **a**, The ds region of the RNA substrate is 17 bp long with overhanging labelled 3' ss on the left, 32 nucleotides in length. The 3' ss right end is 133, and the ss 5' end, 145 nucleotides long. Lane 1, heat-denatured substrate; lanes 2–8, native substrate, incubated with the indicated amounts of p68 under the RNA helicase reaction conditions: lane 2, without p68; lane 3, without ATP; lane 7, without salt; lane 8, preincubated with 5 μ g Pab204 before the addition of the substrate. **b**, 162 bp ds RNA with 32 3' ss nucleotides labelled and the second ss 3' extension 239-nucleotides-long. Lane 1, heat-denatured substrate; lane 2, incubated without p68; lane 3, incubated with 3 ng p68.

METHODS. ssRNA, used to obtain the partial dsRNA substrates, were prepared by run-off transcription, using of appropriate, linearized pGEM3 (Promega) clones (for construction of the clones used, see ref. 16) in opposite orientations to the SP6 and T7 promoters. Transcription assays were run as recommended by the supplier, using [α - 32 P]GTP as the radioactive nucleotide where necessary. Annealing of partially complementary transcripts was performed with a two-fold excess of the unlabelled over the labelled strand in 80% formamide; 0.4 M NaCl; 1 mM EDTA; 40 mM PIPES buffer at pH 6.4. The RNA mixture was heated for



15 min at 80 °C and further incubated for 3 h at 50 °C. The RNA unwinding function of p68 was assayed in 40 μ l volumes containing 2 ng radioactively labelled RNA substrate and the indicated amounts of p68 in a reaction buffer composed of 30 mM Tris-HCl, pH 7.5; 100 mM NaCl; 8 mM MgCl₂; 15 mM dithiothreitol; 30 μ g ml⁻¹ bovine serum albumin; 0.5 unit⁻¹ μ l RNasin and 2 mM ATP. After incubation for 60 min at 37 °C reactions were stopped by addition of 0.1 volume of 3% SDS, 150 mM EDTA.

and II²¹ (Fig. 4 and data not shown). Because it occurred in the presence of several unrelated DNA helicases, it is unlikely that the inhibition results from specific protein-protein interactions. It is not known, however, whether inhibition of DNA unwinding is a physiological function of p68 or represents a residual activity of a DNA helicase, some of whose essential subunit(s) or structure are lost during purification.

We also purified p68 by an immunoaffinity procedure²², using the Pab204. We found DNA- and RNA-binding properties identical to those reported above, whereas the ATPase activity was strongly reduced. Additionally, we could not detect RNA-unwinding activity for the immunoaffinity-purified protein, although inhibition of the DNA helicases was still observed (data not shown). Significantly, the ATPase activity of T antigen eluted from the same antibody column was also inactivated, probably because of the drastic elution conditions (pH 13 buffer).

The p68 protein, including its Pab204 binding site, seems to occur in all dividing eukaryotic cells from the yeast *Saccharo-*

myces cerevisiae (unpublished observations) to man¹, and this evolutionary conservation may reflect a basic function. Besides its amino-acid sequence homology with eIF-4A, it is now clear that both proteins have an RNA-dependent ATPase activity and, moreover, that both proteins have RNA helicase activities that are salt-dependent. Whereas p68 is located in the nucleus, eIF-4A is a cytoplasmic protein involved in the initiation of translation^{13,14} and the two proteins certainly have different functions *in vivo*. Nevertheless, RNA unwinding is most probably an essential step in many processes, including translation and splicing. Splicing of mitochondrial pre-messenger RNA, for instance, has been shown to be dependent on a nuclear protein (pMSS116) that is also classed with the putative helicase family⁷. Interestingly, splicing of nuclear pre-mRNA is accompanied by transient base pairing between U4 and U6 small nuclear RNAs^{23,24}. It is tempting to speculate that p68 may provide the RNA helicase activity for this process and therefore be directly involved in the formation of mature mRNA. □

Received 23 February; accepted 20 April 1989.

1. Lane, D. P. & Hoeffler, W. K. *Nature* **288**, 167–170 (1980).
2. Ford, M. J., Anton, I. A. & Lane, D. P. *Nature* **332**, 736–738 (1988).
3. Nielsen, P. J., McMasters, K. G. & Trachsel, H. *Nucleic Acids Res.* **13**, 6867–6880 (1985).
4. Nielsen, P. J. & Trachsel, H. *EMBO J.* **7**, 2097 (1988).
5. Lasko, P. F. & Ashburner, M. *Nature* **335**, 611–617 (1988).
6. Nishi, K. *et al.* *Nature* **336**, 496–498 (1988).
7. Seraphin, B. *et al.* *Nature* **337**, 84–87 (1989).
8. Linder, P. & Slonimski, P. P. *Proc. natn. Acad. Sci. U.S.A.*, submitted.
9. Leroy, P. *et al.* *Cell*, submitted.
10. Hodgman, T. C. *Nature* **333**, 22–23: 578 (Erratum) (1988).
11. Lane, D. P. *Nature* **334**, 478 (1988).
12. Linder, P. *et al.* *Nature* **337**, 121–122 (1989).
13. Ray, B. K. *et al.* *J. biol. Chem.* **260**, 7651–7658 (1985).
14. Abramson, R. D. *et al.* *J. biol. Chem.* **262**, 3826–3832 (1987).
15. Stahl, H. & Knippers, R. *Biochim. Biophys. Acta* **910**, 1–10 (1987).
16. Scheffner, M., Knippers, R. & Stahl, H. *Cell*, in the press.
17. Rhoads, R. E. *Trends biochem. Sci.* **13**, 52–56 (1988).
18. Grifo, J. A. *et al.* *J. biol. Chem.* **259**, 8648–8654 (1984).
19. Stahl, H. *et al.* *EMBO J.* **5**, 1939–1944 (1986).
20. Wiekowski, M., Schwarz, M. & Stahl, H. *J. biol. Chem.* **263**, 436–442 (1988).
21. Geider, K. & Hoffmann-Berling, H. A. *Rev. Biochem.* **50**, 233–260 (1981).
22. Simanis, V. & Lane, D. P. *Virology* **144**, 88–100 (1985).
23. Lamond, A. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **85**, 411–415 (1988).
24. Brow, D. & Guthrie, C. *Nature* **334**, 213–218 (1988).

ACKNOWLEDGEMENTS. We thank R. Knippers for continuing support and for reviewing the manuscript, T. Kapitza for technical assistance and G. Sweet for reading the manuscript. We also thank D. P. Lane for hybridoma cells producing Pab204 and H. Hoffmann-Berling for gifts of bacterial helicases I and II. This work was supported by the Deutsche Forschungsgemeinschaft.

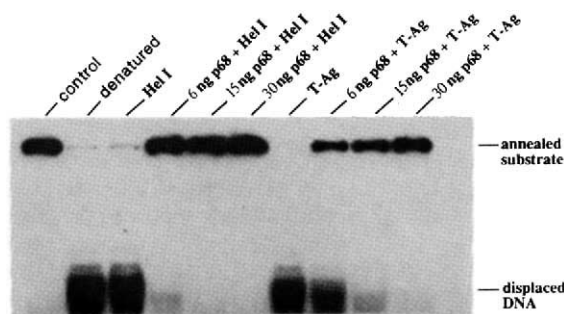


FIG. 4 Inhibition of DNA unwinding by p68. The partially ds DNA substrate (ss M13 DNA containing a 22 bp ds region) and DNA unwinding reactions were exactly as described before^{18,19}. Unwinding products were analysed by SDS-PAGE and autoradiography. Labelling of lanes: Control, incubation without DNA helicase; denatured, heat-denatured substrate DNA; Hel I, incubation with the *E. coli* DNA helicase I (90 ng); T-Ag, incubation with T antigen (270 ng). Where indicated, assay mixtures were preincubated with the respective amount of p68 for 2 min before the addition of the helicase.